Submitted exclusively to the London Mathematical Society
DOI: 10.1112/S0000000000000000

A GENERALIZATION OF VINOGRADOV'S MEAN VALUE THEOREM

SCOTT T. PARSELL

Abstract

We obtain new upper bounds for the number of integral solutions of a complete system of symmetric equations, which may be viewed as a multi-dimensional version of the system considered in Vinogradov's mean value theorem. We then use these bounds to obtain Weyl-type estimates for an associated exponential sum in several variables. Finally, we apply the Hardy-Littlewood method to obtain asymptotic formulas for the number of solutions of the Vinogradov-type system and also of a related system connected to the problem of finding linear spaces on hypersurfaces.

1. *Introduction*

To motivate the topic of this paper, we consider the problem of demonstrating that there exist many rational linear spaces of a given dimension lying on the hypersurface defined by

$$c_1z_1^k + \cdots + c_sz_s^k = 0. \tag{1.1}$$

General results concerning the existence of such spaces are available from work of Brauer [**4**] and Birch [**3**], and estimates for the density of rational lines on (1.1) have been considered in recent work of the author (see [**6**] and [**7**]). A linear space of projective dimension $d-1$ is determined by choosing linearly independent vectors $\mathbf{x}_1, \ldots, \mathbf{x}_d \in \mathbb{Z}^s$. Moreover, the space

$$\mathcal{L}(\mathbf{x}_1, \ldots, \mathbf{x}_d) = \{t_1\mathbf{x}_1 + \cdots + t_d\mathbf{x}_d : t_1, \ldots, t_d \in \mathbb{Q}\}$$

is contained in the hypersurface defined by (1.1) if and only if $\mathbf{x}_1, \ldots, \mathbf{x}_d$ satisfy the system of equations

$$c_1x_{11}^{i_1} \cdots x_{d1}^{i_d} + \cdots + c_sx_{1s}^{i_1} \cdots x_{ds}^{i_d} = 0 \qquad (i_1 + \cdots + i_d = k). \tag{1.2}$$

This is easily seen by substituting into (1.1) and using the multinomial theorem to collect the coefficients of $t_1^{i_1} \cdots t_d^{i_d}$ for each d-tuple $(i_1, \ldots, i_d)$ satisfying $i_1+\cdots+i_d = k$. We shall frequently abbreviate a monomial of the shape $x_1^{i_1} \cdots x_d^{i_d}$ by $\mathbf{x}^\mathbf{i}$. 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 the auxiliary symmetric system

$$\mathbf{x}_1^\mathbf{i} + \cdots + \mathbf{x}_s^\mathbf{i} = \mathbf{y}_1^\mathbf{i} + \cdots + \mathbf{y}_s^\mathbf{i} \qquad (i_1 + \cdots + i_d = k) \tag{1.3}$$

lying in a given box. Our strategy for obtaining such estimates is similar to that encountered in the application of Vinogradov's mean value theorem to Waring's problem. Specifically, we consider the augmented system

$$\mathbf{x}_1^\mathbf{i} + \cdots + \mathbf{x}_s^\mathbf{i} = \mathbf{y}_1^\mathbf{i} + \cdots + \mathbf{y}_s^\mathbf{i} \qquad (1 \le i_1 + \cdots + i_d \le k), \tag{1.4}$$

2000 *Mathematics Subject Classification* 11D45, 11D72, 11L07, 11P55.

Supported in part by a National Science Foundation Postdoctoral Fellowship (DMS-0102068)

where the number of equations here is

$$r = \binom{k+d}{d} - 1. \tag{1.5}$$

Note that the classical version of Vinogradov's mean value theorem (see for example [**10**]) is concerned with the system (1.4) in the case $d = 1$. In this case, the sharpest available results are due to Wooley [**12**]. The presence of the equations of lower degree facilitates the application of a p-adic iteration method, in which repeated use of the binomial theorem makes it essential to consider such equations together with those of degree k.

In order to count solutions of (1.4), we need to analyze the exponential sum

$$f(\boldsymbol{\alpha}) = f(\boldsymbol{\alpha}; P) = \sum_{\mathbf{x}\in[1,P]^d} e\left(\sum_{1\le|\mathbf{i}|\le k} \alpha_{\mathbf{i}}\mathbf{x}^{\mathbf{i}}\right), \tag{1.6}$$

where we have written $e(y) = e^{2\pi i y}$ and $|\mathbf{i}| = i_1 + \cdots + i_d$. Here and throughout, we suppose that P is sufficiently large in terms of s, k, and d. Furthermore, we take d to be fixed and suppose that k is sufficiently large in terms of d. Let $J_{s,k,d}(P)$ denote the number of solutions of the system (1.4) with $\mathbf{x}_m, \mathbf{y}_m \in [1, P]^d \cap \mathbb{Z}^d$. Then by orthogonality we have

$$J_{s,k,d}(P) = \int_{\mathbb{T}^r} |f(\boldsymbol{\alpha})|^{2s}\, d\boldsymbol{\alpha}, \tag{1.7}$$

where $\mathbb{T}^r$ denotes the r-dimensional unit cube. Before considering upper bounds for $J_{s,k,d}(P)$, it is useful to derive an elementary lower bound. Let $J_{s,k,d}(P;\mathbf{h})$ denote the number of solutions of the system

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

with $\mathbf{x}_m, \mathbf{y}_m \in [1, P]^d \cap \mathbb{Z}^d$, and observe that

$$J_{s,k,d}(P;\mathbf{h}) = \int_{\mathbb{T}^r} |f(\boldsymbol{\alpha})|^{2s} e(-\boldsymbol{\alpha}\cdot\mathbf{h})\, d\boldsymbol{\alpha} \le J_{s,k,d}(P). \tag{1.8}$$

Thus, by summing over all values of $\mathbf{h}$ for which $J_{s,k,d}(P;\mathbf{h})$ is nonzero, we find that

$$J_{s,k,d}(P) \gg P^{2sd-K}, \tag{1.9}$$

where

$$K = \sum_{l=1}^{k} l\binom{l+d-1}{l} \tag{1.10}$$

is the sum of the degrees of the equations in (1.4). By considering diagonal solutions, one also obtains the lower bound $J_{s,k,d}(P) \gg P^{sd}$, but the expression in (1.9) dominates whenever $s > K/d$. Moreover, an informal probabilistic argument suggests that P^{2sd-K} represents the true order of magnitude. Thus we aim to establish estimates of the shape

$$J_{s,k,d}(P) \ll P^{2sd-K+\Delta_s}, \tag{1.11}$$

where $\Delta_s = \Delta_{s,k,d}$ is small whenever s is sufficiently large in terms of k and d. Whenever an estimate of the form (1.11) holds, we say that Δ_s is an admissible exponent for (s, k, d). Our main theorem is the following.

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

$$s_0 = rk(\tfrac{1}{2}\log k - \log\log k). \tag{1.12}$$

Then the estimate (1.11) *holds with*

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

Somewhat more refined (and complicated) conclusions are given in Theorems 4.2 and 4.5 below, but the simplified version given above is sufficient for most applications. We also note that Arkhipov, Chubarikov, and Karatsuba [**1**] have obtained results of this type for related systems in which our condition $1 \le |\mathbf{i}| \le k$ is weakened to $\mathbf{i} \in [0,k]^d$. While estimates for $J_{s,k,d}(P)$ can be derived from their results, it is clear that admissible exponents decaying roughly like $rke^{-s/rk}$ are the best that could be extracted from their methods. The superior decay achieved in Theorem 1.1 results from a repeated efficient differencing approach first devised by Wooley [**12**].

Standard methods exist for translating mean value estimates such as those given in Theorem 1.1 into Weyl-type estimates for the exponential sum $f(\boldsymbol{\alpha})$, and we state one such theorem below. When $\mathbf{a}$ is a vector in $\mathbb{Z}^n$, we find it useful to write $(q, \mathbf{a})$ for $\gcd(q, a_1, \dots, a_n)$.

Theorem 1.2. *Suppose that k is sufficiently large in terms of d and that $|f(\boldsymbol{\alpha})| \ge P^{d-\sigma+\varepsilon}$ for some $\varepsilon > 0$, where $\sigma^{-1} \ge \frac{8}{3}rk\log rk$. Then there exist integers $a_{\mathbf{j}}$ and q, with $(q,\mathbf{a}) = 1$, satisfying*

$$1 \le q \le P^{k\sigma} \quad \text{and} \quad |q\alpha_{\mathbf{j}} - a_{\mathbf{j}}| \le P^{k\sigma-|\mathbf{j}|} \qquad (1 \le |\mathbf{j}| \le k).$$

Some amount of technical effort is required to prove this, and in §5 we need to establish some auxiliary results of this type (see Theorems 5.1 and 5.2), which may be of interest in their own right for certain applications. A slightly sharper form of Theorem 1.2 is actually given in Theorem 5.5, but the former suffices for our purposes. For smaller k, one may be able to obtain results of this nature by a Weyl differencing argument (see [**6**] for the case $k = 3$ and $d = 2$).

By applying Theorems 1.1 and 1.2 within the Hardy-Littlewood method, one can show that $\Delta_s = 0$ is admissible in (1.11) when s is sufficiently large in terms of k and d. Furthermore, the method yields an asymptotic formula for $J_{s,k,d}(P)$.

Theorem 1.3. *Suppose that k is sufficiently large in terms of d, and write*

$$s_1 = rk(\tfrac{2}{3}\log r + \tfrac{1}{2}\log k + \log\log k + 2d + 4). \tag{1.13}$$

There are positive constants $C = C(s,k,d)$ and $\delta = \delta(k,d)$ such that, whenever $s \ge s_1$, one has

$$J_{s,k,d}(P) = (C + O(P^{-\delta}))P^{2sd-K}.$$

When $d = 1$, the lower bound on s can be improved somewhat. Specifically, the coefficient $2/3$ in the $\log r$ term can be replaced by $1/2$ (see Wooley [**15**], Theorem 3), so that $s_1 \sim k^2 \log k$. The reason for this, roughly speaking, is that the number of linearly independent monomials $\mathbf{x}^{\mathbf{i}}$ with $|\mathbf{i}| \le k - j$ differs from r by essentially

jk^{d-1}. The resulting loss of congruence data in the j-fold repeated differencing algorithm (measured by the sum of the degrees of the equations that must be ignored) therefore behaves roughly like jk^d in general. When $d = 1$, however, the loss is only about j^2, since each equation that is removed has degree at most j. Typically one takes j to be a power of $\log k$, so it turns out that an unusually large amount of information is retained in the $d = 1$ case as compared to the situation when $d \geq 2$.

We now indicate how to use estimates of the type (1.11) to obtain bounds for the number of solutions of (1.3), which are relevant to counting linear spaces on hypersurfaces. If we let $I_{s,k,d}(P)$ denote the number of solutions of the system (1.3) with $\mathbf{x}_m, \mathbf{y}_m \in [1, P]^d \cap \mathbb{Z}^d$, then one has

$$I_{s,k,d}(P) = \sum_{\mathbf{h}} J_{s,k,d}(P; \mathbf{h}),$$

where the summation is over all vectors $\mathbf{h} \in \mathbb{Z}^r$ with $h_{\mathbf{i}} = 0$ when $|\mathbf{i}| = k$. The number of choices of $\mathbf{h}$ for which $J_{s,k,d}(P; \mathbf{h}) \neq 0$ is $O(P^{K-L})$, where we have written

$$L = k\binom{k+d-1}{k} \tag{1.14}$$

for the sum of the degrees of the equations in (1.3). We therefore see from (1.8) that the estimate (1.11) yields

$$I_{s,k,d}(P) \ll J_{s,k,d}(P)P^{K-L} \ll P^{2sd-L+\Delta_s}. \tag{1.15}$$

Moreover, by imitating the argument leading to (1.9), one finds that $I_{s,k,d}(P) \gg P^{2sd-L}$, so in each case Δ_s measures the difference between the exponent in our attainable bound and the best possible exponent. It is conceivable that a more sophisticated strategy along the lines of Ford [**5**] could be applied to relate $I_{s,k,d}(P)$ to $J_{s,k,d}(P)$, but we do not pursue this here.

Estimates of the shape (1.15) enable one to establish an asymptotic formula for the number of solutions of the system (1.2) lying in a given box, provided that s is sufficiently large in terms of k and d and that certain local solubility conditions are satisfied. Let $N_{s,k,d}(P)$ denote the number of solutions of the system (1.2) with $x_{lm} \in [-P, P] \cap \mathbb{Z}$. The proof of the following theorem follows essentially the same pattern as the proof of Theorem 1.3.

Theorem 1.4. *Suppose that k is sufficiently large in terms of d and that $s \geq 2s_1$, where s_1 is as in* (1.13). *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 $\nu = \nu(k, d)$ such that*

$$N_{s,k,d}(P) = (\mathcal{C} + O(P^{-\nu}))P^{sd-L}.$$

In particular, this establishes the existence of many rational linear spaces of projective dimension $d-1$ on the hypersurface (1.1), provided that $s \geq 2s_1$ and that the appropriate local solubility conditions are met. Roughly speaking, the theorem counts linear spaces up to a given height, weighted according to the number of integral bases.

We sketch a proof of Theorem 1.4 towards the end of the paper, but our main focus here is on the estimates of Theorems 1.1–1.3 for $J_{s,k,d}(P)$ and the associated exponential sum $f(\boldsymbol{\alpha})$. In a future paper, we plan to investigate estimates for the number of solutions of (1.2) in greater detail. In particular, if one only desires an asymptotic lower bound for the number of solutions, then one can restrict to solutions in R-smooth numbers, where R is a small power of P. In this case, the system (1.3) can be considered directly by a variant of the Vaughan-Wooley iterative method, and as a result the number of variables needed is reduced by roughly a factor of k, just as in the situation of Waring's problem. Thus, while something on the order of $k^{d+1}\log k$ variables is required to prove the asymptotic formula, one should be able to establish asymptotic lower bounds with only on the order of $k^d \log k$ variables. This latter expectation has already been established by the author [**7**] in the case $d = 2$ with a leading coefficient of 14/3. For smaller k, one can perform more precise analyses along the lines of [**6**] to obtain explicit numerical bounds on the number of variables required.

2. *Preliminary observations*

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 (1.4). 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 introduce the notation

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

for the number of equations in (1.4) with $\mathbf{i} \succ \mathbf{j}_1$, where we have written $\mathbf{j}_1$ for the vector $(j, 0, \dots, 0)$. Observe that r_j is also the number of distinct monomials $\mathbf{x}^{\mathbf{i}}$ with $1 \le |\mathbf{i}| \le k - j$. We further write

$$K_j = \sum_{l=j+1}^{k} l \binom{l-j+d-1}{l-j} \tag{2.2}$$

for the sum of the degrees of the equations in (1.4) with $\mathbf{i} \succ \mathbf{j}_1$. In particular, we recall from (1.10) that $K_0 = K$. Before proceeding, we find it useful to record a closed formula for K_j.

Lemma 2.1. *For $0 \le j \le k$, one has*

$$K_j = \frac{dk+j}{d+1}\binom{k-j+d}{d} - j.$$

Proof. We first establish the formula for $j = 0$. We have

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

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

$$K = d\binom{k+d}{k-1} = d\binom{k+d}{d+1} = \frac{dk}{d+1}\binom{k+d}{d}, \tag{2.3}$$

as required. To handle K_j, we first re-index the sum (2.2) to get

$$K_j = \sum_{l=1}^{k-j} (l+j) \binom{l+d-1}{l} = K[k-j] + jr[k-j],$$

where $K[k-j]$ and $r[k-j]$ denote the parameters K and r with k replaced by $k-j$. Thus by applying (1.5) and (2.3) we obtain

$$K_j = \frac{d(k-j)}{d+1} \binom{k-j+d}{d} + j \left[\binom{k-j+d}{d} - 1 \right],$$

and the lemma now follows easily. □

Next, we let $\mathcal{B}_{p,j}(\mathbf{f};\mathbf{u})$ denote the number of solutions $\mathbf{x}$ modulo p^k of the system

$$f_{\mathbf{i}}(\mathbf{x}) \equiv u_{\mathbf{i}} \pmod{p^{|\mathbf{i}|}} \qquad (\mathbf{i} \succ \mathbf{j}_1)$$

for which the rank of the Jacobian matrix $(\partial f_{\mathbf{i}}/\partial x_l)$ modulo p is r_j.

Lemma 2.2. *Let r_j and K_j be as in* (2.1) *and* (2.2), *and let p be a prime. If each $f_{\mathbf{i}}$ is a polynomial in t variables with integer coefficients and $t \geq r_j$, then one has*

$$\operatorname{card} \mathcal{B}_{p,j}(\mathbf{f};\mathbf{u}) \ll p^{kt-K_j},$$

where the implicit constant depends at most on the degrees of the $f_{\mathbf{i}}$.

Proof. We start by choosing integers $a_{\mathbf{i}} \equiv u_{\mathbf{i}} \pmod{p^{|\mathbf{i}|}}$ with $1 \leq a_{\mathbf{i}} \leq p^k$ for each $\mathbf{i}$ with $\mathbf{i} \succ \mathbf{j}_1$. It follows from the main theorem of Wooley [**14**] that the number of non-singular solutions of the system of congruences

$$f_{\mathbf{i}}(\mathbf{x}) \equiv a_{\mathbf{i}} \pmod{p^k} \qquad (\mathbf{i} \succ \mathbf{j}_1)$$

is $O(p^{k(t-r_j)})$ for each choice of $\mathbf{a}$. Now the number of choices for $\mathbf{a}$ is p^ω, where

$$\omega = \sum_{\mathbf{i} \succ \mathbf{j}_1} (k - |\mathbf{i}|) = kr_j - K_j,$$

and thus $\operatorname{card} \mathcal{B}_{p,j}(\mathbf{f};\mathbf{u}) \ll p^{kr_j - K_j} \cdot p^{k(t-r_j)} = p^{kt-K_j}$. □

The following result, based on the binomial theorem, enables our p-adic iteration by transforming a system in which certain variables are classified according to residue class modulo p to one in which a power of p divides both sides of each equation in the system. This facilitates the introduction of a strong congruence condition on the remaining variables, and it is here that we require the presence of the equations of lower degree in (1.4). The method cannot be applied directly to the system (1.3) that one wants to consider for the application to linear spaces on hypersurfaces.

In what follows, when $\mathbf{x}$, $\mathbf{a}$, and $\mathbf{i}$ are d-dimensional vectors and p is a scalar, we adopt the notation $(p\mathbf{x}+\mathbf{a})^{\mathbf{i}} = (px_1+a_1)^{i_1} \cdots (px_d+a_d)^{i_d}$. We also let $\Psi_{\mathbf{i}}(\mathbf{z})$ denote any function of d variables and let $\eta_1, \dots, \eta_n$ denote any real numbers.

Lemma 2.3. *Every solution* $(\mathbf{z}, \mathbf{w}, \mathbf{x}, \mathbf{y})$ *of the system*

$$\sum_{n=1}^{r} \eta_n(\Psi_{\mathbf{i}}(\mathbf{z}_n) - \Psi_{\mathbf{i}}(\mathbf{w}_n)) = \sum_{m=1}^{s} ((p\mathbf{x}_m+\mathbf{a})^{\mathbf{i}} - (p\mathbf{y}_m+\mathbf{a})^{\mathbf{i}}) \quad (1 \leq |\mathbf{i}| \leq k) \tag{2.4}$$

is a solution of the system

$$\sum_{n=1}^{r} \eta_n(\Phi_{\mathbf{i}}(\mathbf{z}_n) - \Phi_{\mathbf{i}}(\mathbf{w}_n)) = p^{|\mathbf{i}|} \sum_{m=1}^{s} (\mathbf{x}_m^{\mathbf{i}} - \mathbf{y}_m^{\mathbf{i}}) \qquad (1 \le |\mathbf{i}| \le k), \tag{2.5}$$

where

$$\Phi_{\mathbf{i}}(\mathbf{z}) = \sum_{l_1=0}^{i_1} \cdots \sum_{l_d=0}^{i_d} \binom{i_1}{l_1} \cdots \binom{i_d}{l_d} (-\mathbf{a})^{\mathbf{i}-\mathbf{l}} \Psi_{\mathbf{l}}(\mathbf{z}). \tag{2.6}$$

Conversely, every solution of (2.5) *is a solution of* (2.4).

Proof. Suppose that $(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w})$ satisfies (2.4). By the binomial theorem, we have

$$(px)^i = \sum_{l=0}^{i} \binom{i}{l} (px+a)^l (-a)^{i-l},$$

where we adopt the convention that $0^0 = 1$. Thus the right-hand side of (2.5) can be expressed as

$$\begin{aligned}
&\sum_{m=1}^{s} \left(\prod_{j=1}^{d} \sum_{l=0}^{i_j} \binom{i_j}{l} (px_{mj}+a_j)^l (-a_j)^{i_j-l} - \prod_{j=1}^{d} \sum_{l=0}^{i_j} \binom{i_j}{l} (py_{mj}+a_j)^l (-a_j)^{i_j-l} \right) \\
&= \sum_{l_1=0}^{i_1} \cdots \sum_{l_d=0}^{i_d} \binom{i_1}{l_1} \cdots \binom{i_d}{l_d} (-\mathbf{a})^{\mathbf{i}-\mathbf{l}} \sum_{m=1}^{s} [(p\mathbf{x}_m + \mathbf{a})^{\mathbf{l}} - (p\mathbf{y}_m + \mathbf{a})^{\mathbf{l}}] \\
&= \sum_{n=1}^{r} \eta_n (\Phi_{\mathbf{i}}(\mathbf{z}_n) - \Phi_{\mathbf{i}}(\mathbf{w}_n)),
\end{aligned}$$

on substituting (2.4) and recalling (2.6). Conversely, suppose that $(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w})$ satisfies (2.5). Then on applying the binomial theorem, we find that the right-hand side of (2.4) is given by

$$\begin{aligned}
S_{\mathbf{i}} &= \sum_{m=1}^{s} \left(\prod_{j=1}^{d} \sum_{l=0}^{i_j} \binom{i_j}{l} (px_{mj})^l a_j^{i_j-l} - \prod_{j=1}^{d} \sum_{l=0}^{i_j} \binom{i_j}{l} (py_{mj})^l a_j^{i_j-l} \right) \\
&= \sum_{l_1=0}^{i_1} \cdots \sum_{l_d=0}^{i_d} \binom{i_1}{l_1} \cdots \binom{i_d}{l_d} \mathbf{a}^{\mathbf{i}-\mathbf{l}} \sum_{n=1}^{r} \eta_n (\Phi_{\mathbf{l}}(\mathbf{z}_n) - \Phi_{\mathbf{l}}(\mathbf{w}_n)).
\end{aligned}$$

On substituting (2.6) and interchanging the order of summation, we obtain

$$S_{\mathbf{i}} = \sum_{j_1=0}^{i_1} \cdots \sum_{j_d=0}^{i_d} (-1)^{-|\mathbf{j}|} \mathbf{a}^{\mathbf{i}-\mathbf{j}} \Theta(\mathbf{i}, \mathbf{j}) \sum_{n=1}^{r} \eta_n (\Psi_{\mathbf{j}}(\mathbf{z}_n) - \Psi_{\mathbf{j}}(\mathbf{w}_n)),$$

where

$$\Theta(\mathbf{i}, \mathbf{j}) = \prod_{t=1}^{d} \sum_{l=j_t}^{i_t} (-1)^l \binom{i_t}{l} \binom{l}{j_t}.$$

It now suffices to note that

$$\sum_{l=j}^{i} (-1)^l \binom{i}{l} \binom{l}{j} = \binom{i}{j} \sum_{l=j}^{i} (-1)^l \binom{i-j}{i-l} = (-1)^i \binom{i}{j} \sum_{l=0}^{i-j} \binom{i-j}{l} (-1)^l,$$

which equals $(-1)^j$ if $i = j$ and is zero otherwise. Therefore $\Theta(\mathbf{i}, \mathbf{j}) = (-1)^{|\mathbf{j}|}$ if $\mathbf{j} = \mathbf{i}$ and $\Theta(\mathbf{i}, \mathbf{j}) = 0$ otherwise, whence the only contribution to the expression for $S_{\mathbf{i}}$ comes from the terms with $\mathbf{j} = \mathbf{i}$. □

Next we need to say something about the types of system we will be working with in our applications. The following definition covers the systems we need to study.

DEFINITION 2.4. *We say that the system of polynomials* $(\mathbf{\Psi})$ *is of type* (j, P, A) *if the following conditions are satisfied.*

(1) *The system consists of polynomials* $\Psi_{\mathbf{i}} \in \mathbb{Z}[z_1, \ldots, z_d]$*, indexed by the vectors* $\mathbf{i}$ *satisfying* $1 \le |\mathbf{i}| \le k$.

(2) *The polynomial* $\Psi_{\mathbf{i}}$ *has degree* $|\mathbf{i}| - j$ *when* $|\mathbf{i}| \ge j$ *and is identically zero otherwise.*

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

(4) *For each* $\mathbf{i}$ *with* $\mathbf{i} \succ (j, 0, \ldots, 0)$*, the polynomial* $\Psi_{\mathbf{i}}$ *contains a term of degree* $|\mathbf{i}| - j$ *that does not appear explicitly in any of the* $\Psi_{\mathbf{i}'}$ *with* $|\mathbf{i}'| = |\mathbf{i}|$ *and* $\mathbf{i}' \succ \mathbf{i}$.

Condition (4) may be viewed as a sort of linear independence requirement and will be important in estimating the number of singular solutions of our systems of congruences. We also mention that if the system $(\mathbf{\Psi})$ is of type (j, P, A), then the system $(\mathbf{\Phi})$ defined by (2.6) is also of type (j, P, A), since the terms of highest degree in $\Psi_{\mathbf{i}}$ and $\Phi_{\mathbf{i}}$ are identical.

3. *The efficient differencing apparatus*

Fix k and d, let θ be a parameter with $0 < \theta \le 1/k$, and suppose that $(\mathbf{\Psi})$ is a system of type (j, P, A). Further, write $\mathbf{j}_1 = (j, 0, \ldots, 0)$. Then all the coefficients of the terms of highest degree in each of the polynomials

$$\frac{\partial \Psi_{\mathbf{i}}}{\partial z_l}(\mathbf{z}) \qquad (\mathbf{i} \succ \mathbf{j}_1,\ 1 \le l \le d)$$

are bounded in absolute value by kAP^k, so the number of prime divisors p of a given non-zero coefficient with $p > P^\theta$ is bounded in terms of k, A, and θ. Furthermore, the total number of coefficients under consideration is bounded in terms of k and d, so the total number of prime divisors of all these coefficients is bounded by a constant $c = c(k, d, A, \theta)$. We let $\mathcal{P}(\theta)$ denote the set consisting of the smallest $c + [1/\theta]$ primes exceeding P^θ. Clearly, if P is sufficiently large, then the Prime Number Theorem ensures that $P^\theta < p < 2P^\theta$ for all $p \in \mathcal{P}(\theta)$.

For simplicity, we often write $J_s(P)$ for $J_{s,k,d}(P)$. Our goal in this section is to develop an iterative method for bounding $J_s(P)$ as s increases, and it is convenient to increase s to $s + r$, where r is as in (1.5), at each stage of the iteration. Thus we let $K_s(P, Q; \mathbf{\Psi})$ denote the number of solutions of the system

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

with $1 \le z_{nl}, w_{nl} \le P$ and $1 \le x_{ml}, y_{ml} \le Q$. We also write $\mathrm{Jac}(\mathbf{\Psi}; \mathbf{z}, \mathbf{w})$ for the $r_j \times 2rd$ Jacobian matrix formed with the polynomials on the left-hand side for which $\mathbf{i} \succ \mathbf{j}_1$. Further, we let $L_s(P,Q,\theta,p;\mathbf{\Psi})$ denote the number of solutions of the system

$$\sum_{n=1}^{r}(\Psi_{\mathbf{i}}(\mathbf{z}_n)-\Psi_{\mathbf{i}}(\mathbf{w}_n)) = p^{|\mathbf{i}|}\sum_{m=1}^{s}(\mathbf{u}_m^{\mathbf{i}}-\mathbf{v}_m^{\mathbf{i}}) \qquad (1 \le |\mathbf{i}| \le k) \tag{3.2}$$

with $\mathbf{z}$ and $\mathbf{w}$ as above, with $1 \le u_{ml}, v_{ml} \le QP^{-\theta}$, and with $z_{nl} \equiv w_{nl} \pmod{p^k}$. Finally, we write

$$L_s(P,Q,\theta;\mathbf{\Psi}) = \max_{p\in\mathcal{P}(\theta)} L_s(P,Q,\theta,p;\mathbf{\Psi}).$$

We are now ready to state our fundamental lemma. In what follows, we find it convenient to write

$$q_j = \binom{j+d}{d} - 1$$

to denote the number of equations in (1.4) of total degree at most j.

Lemma 3.1. *Suppose that $s \ge 2q_j - 1$, that $P^\theta \le Q \le P$, and that $(\mathbf{\Psi})$ is a system of type (j,P,A) for some constant $A = A(k,d)$. Then there is a system $(\mathbf{\Phi})$ of type (j,P,A), given by (2.6), such that*

$$K_s(P,Q;\mathbf{\Psi}) \ll P^{2rd-(1-\theta)(r+1)}J_s(Q) + P^{\theta(2sd+\omega(k,j,d))}L_s(P,Q,\theta;\mathbf{\Phi}),$$

where

$$\omega(k,j,d) = krd - K_j - q_j.$$

Proof. First of all, let S_1 denote the number of solutions of (3.1) for which the rank modulo p of $\mathrm{Jac}(\mathbf{\Psi};\mathbf{z},\mathbf{w})$ is less than r_j for all primes $p \in \mathcal{P}(\theta)$. Consider a choice of $\mathbf{z}$ and $\mathbf{w}$ counted by S_1. By construction, there exist distinct primes $p_1,\dots,p_t \in \mathcal{P}(\theta)$, where $t = [1/\theta]$, none of which divides any coefficient of a term of maximal degree in any of the polynomials $\partial\Psi_{\mathbf{i}}/\partial z_l$. Let p denote any one of the primes $p_1,\dots,p_t$. If the rank modulo p of $\mathrm{Jac}(\mathbf{\Psi};\mathbf{z},\mathbf{w})$ is less than r_j, then there exists a non-trivial linear relation over $\mathbb{F}_p$ among the rows of this matrix. That is, there exist $c_{\mathbf{i}} \in \mathbb{F}_p$, not all zero, such that

$$\sum_{\mathbf{i}\succ\mathbf{j}_1} c_{\mathbf{i}}\frac{\partial\Psi_{\mathbf{i}}}{\partial z_l}(\mathbf{z}) \equiv 0 \pmod{p} \tag{3.3}$$

for $\mathbf{z} = \mathbf{z}_1,\dots,\mathbf{z}_r,\mathbf{w}_1,\dots,\mathbf{w}_r$ and $l = 1,\dots,d$. The number of choices for the coefficients $c_{\mathbf{i}}$ is $O(p^{r_j-1})$, since one of them may be normalized to 1. Now let I denote the largest value of $|\mathbf{i}|$ for which the corresponding $c_{\mathbf{i}}$ is non-zero, and let $\mathbf{i}$ denote the smallest index (in the lexicographic ordering defined above) for which $|\mathbf{i}| = I$ and $c_{\mathbf{i}}$ is non-zero modulo p. By condition (4) of Definition 2.4, there is an l with $1 \le l \le d$ such that $\partial\Psi_{\mathbf{i}}/\partial z_l$ contains a term of degree $I-j-1$ that is not present in any $\partial\Psi_{\mathbf{j}}/\partial z_l$ with $|\mathbf{j}| = I$ and $\mathbf{j} \succ \mathbf{i}$, and this term is nonzero modulo p by the definition of $p_1,\dots,p_t$. Thus, by considering terms of degree $I-j-1$, it follows from the maximality of I that the polynomial on the left-hand side of (3.3) is not identically zero in $\mathbb{F}_p[\mathbf{z}]$. Hence each $\mathbf{z}_n$ and $\mathbf{w}_n$ satisfies a non-trivial polynomial in d variables over the field $\mathbb{F}_p$, so the argument of the proof of Lemma 2 of Wooley [**13**] 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 $c_{\mathbf{i}}$. Thus, by the Chinese Remainder Theorem, the total number of possibilities for $\mathbf{z}$ and $\mathbf{w}$ modulo $p_1 \cdots p_t$ is $\ll (p_1 \cdots p_t)^{2rd-r-1}$. For each such choice, there are trivially at most $(P/(p_1 \cdots p_t))^{2rd}$ choices for $\mathbf{z}$ and $\mathbf{w}$, so it follows from (1.8) that

$$S_1 \ll P^{2rd}(p_1 \cdots p_t)^{-r-1} J_s(Q) \ll P^{2rd-(r+1)(1-\theta)} J_s(Q). \tag{3.4}$$

Now let S_2 be the number of solutions for which the rank modulo p of $\mathrm{Jac}(\boldsymbol{\Psi}; \mathbf{z}, \mathbf{w})$ is r_j for some prime $p \in \mathcal{P}(\theta)$; here p may of course depend on $\mathbf{z}$ and $\mathbf{w}$. Then one has

$$S_2 \le \sum_{p \in \mathcal{P}(\theta)} S_3(p),$$

where $S_3(p)$ is the number of solutions of (3.1) with $\mathrm{Jac}(\boldsymbol{\Psi}; \mathbf{z}, \mathbf{w})$ having rank r_j modulo p. Write

$$G(\boldsymbol{\alpha}; \boldsymbol{\eta}) = \sum_{\mathbf{z} \in [1,P]^{rd}} e\left(\sum_{1 \le |\mathbf{i}| \le k} \alpha_{\mathbf{i}} s_{\mathbf{i}}(\mathbf{z}; \boldsymbol{\eta}) \right),$$

where

$$s_{\mathbf{i}}(\mathbf{z}; \boldsymbol{\eta}) = \eta_1 \Psi_{\mathbf{i}}(\mathbf{z}_1) + \cdots + \eta_r \Psi_{\mathbf{i}}(\mathbf{z}_r),$$

and let $G_p(\boldsymbol{\alpha}; \boldsymbol{\eta})$ denote the same sum, but restricted to those $\mathbf{z}$ for which the $r_j \times rd$ matrix $\mathrm{Jac}(\boldsymbol{\Psi}; \mathbf{z})$ has rank r_j modulo p. After rearranging variables, one finds that

$$S_3(p) \le \sum_{\boldsymbol{\eta} \in \{\pm 1\}^r} \int_{\mathbb{T}^r} G(\boldsymbol{\alpha}; \boldsymbol{\eta}) G_p(-\boldsymbol{\alpha}; \boldsymbol{\eta}) |f(\boldsymbol{\alpha}; Q)|^{2s} \, d\boldsymbol{\alpha},$$

so by applying the Cauchy-Schwarz inequality we get

$$S_3(p) \ll \left(\int_{\mathbb{T}^r} |G(\boldsymbol{\alpha}; \boldsymbol{\eta})|^2 |f(\boldsymbol{\alpha}; Q)|^{2s} \, d\boldsymbol{\alpha} \right)^{1/2} \left(\int_{\mathbb{T}^r} |G_p(\boldsymbol{\alpha}; \boldsymbol{\eta})|^2 |f(\boldsymbol{\alpha}; Q)|^{2s} \, d\boldsymbol{\alpha} \right)^{1/2}$$

for some $\boldsymbol{\eta} \in \{\pm 1\}^r$. It is easy to see that $G(\boldsymbol{\alpha}; \boldsymbol{\eta})$ may be expressed as a product of r exponential sums, each in d variables. It follows by taking complex conjugates that $|G(\boldsymbol{\alpha}; \boldsymbol{\eta})| = |G(\boldsymbol{\alpha}; \mathbf{1})|$ and hence that the integral in the first factor above is equal to $K_s(P, Q; \boldsymbol{\Psi})$. Suppose that $S_2 \ge S_1$. Then on noting that $|\mathcal{P}(\theta)| \ll 1$, we find that

$$K_s(P, Q; \boldsymbol{\Psi}) = S_1 + S_2 \ll \max_{\substack{p \in \mathcal{P}(\theta) \\ \boldsymbol{\eta} \in \{\pm 1\}^r}} S_4(p; \boldsymbol{\eta}), \tag{3.5}$$

where $S_4(p; \boldsymbol{\eta})$ denotes the number of solutions of the system

$$\sum_{n=1}^{r} \eta_n (\Psi_{\mathbf{i}}(\mathbf{z}_n) - \Psi_{\mathbf{i}}(\mathbf{w}_n)) = \sum_{m=1}^{s} (\mathbf{x}_m^{\mathbf{i}} - \mathbf{y}_m^{\mathbf{i}}) \qquad (1 \le |\mathbf{i}| \le k) \tag{3.6}$$

with both $\mathrm{Jac}(\boldsymbol{\Psi}; \mathbf{z})$ and $\mathrm{Jac}(\boldsymbol{\Psi}; \mathbf{w})$ having rank r_j modulo p.

Since $(\boldsymbol{\Psi})$ is of type (j, P, A), we have

$$\sum_{m=1}^{s} (\mathbf{x}_m^{\mathbf{i}} - \mathbf{y}_m^{\mathbf{i}}) = 0 \qquad (1 \le |\mathbf{i}| \le j),$$

so we can classify the solutions counted by $S_4(p)$ according to the common residue classes of $\mathbf{x}_1^{\mathbf{i}} + \cdots + \mathbf{x}_s^{\mathbf{i}}$ and $\mathbf{y}_1^{\mathbf{i}} + \cdots + \mathbf{y}_s^{\mathbf{i}}$ modulo p. Thus we write $\mathcal{B}_p(\mathbf{w})$ for the set of solutions modulo p of the system of congruences

$$\sum_{m=1}^{s} \mathbf{x}_m^{\mathbf{i}} \equiv w_{\mathbf{i}} \pmod p \qquad (1 \le |\mathbf{i}| \le j).$$

The main theorem of Wooley [**14**] shows that the number of non-singular solutions counted by $\mathcal{B}_p(\mathbf{w})$ is $O(p^{sd-q_j})$. Moreover, since $p \in \mathcal{P}(\theta)$, the argument used in connection with the estimation of S_1 above shows that the number of singular solutions is $O(p^{q_j-1+s(d-1)})$. We therefore deduce that

$$\text{card } \mathcal{B}_p(\mathbf{w}) \ll p^{sd-q_j}, \tag{3.7}$$

provided that $s \ge 2q_j - 1$. We now introduce the exponential sum

$$f_p(\boldsymbol{\alpha}; \mathbf{y}) = \sum_{\substack{\mathbf{x} \in [1,Q]^d \\ \mathbf{x} \equiv \mathbf{y} \pmod p}} e\left(\sum_{1 \le |\mathbf{i}| \le k} \alpha_{\mathbf{i}} \mathbf{x}^{\mathbf{i}} \right)$$

and note that

$$S_4(p; \boldsymbol{\eta}) = \int_{\mathbb{T}^r} |G_p(\boldsymbol{\alpha}; \boldsymbol{\eta})|^2 \sum_{\mathbf{w} \in [1,p]^{q_j}} |U_p(\boldsymbol{\alpha}; \mathbf{w})|^2 \, d\boldsymbol{\alpha},$$

where

$$U_p(\boldsymbol{\alpha}; \mathbf{w}) = \sum_{(\mathbf{u}_1, \ldots, \mathbf{u}_s) \in \mathcal{B}_p(\mathbf{w})} f_p(\boldsymbol{\alpha}; \mathbf{u}_1) \cdots f_p(\boldsymbol{\alpha}; \mathbf{u}_s).$$

It follows from Cauchy's inequality and (3.7) that

$$\begin{aligned} |U_p(\boldsymbol{\alpha}; \mathbf{w})|^2 &\ll \text{card } \mathcal{B}_p(\mathbf{w}) \sum_{\mathbf{u} \in \mathcal{B}_p(\mathbf{w})} |f_p(\boldsymbol{\alpha}; \mathbf{u}_1) \cdots f_p(\boldsymbol{\alpha}; \mathbf{u}_s)|^2 \\ &\ll p^{sd-q_j} \sum_{\mathbf{u} \in \mathcal{B}_p(\mathbf{w})} \sum_{i=1}^{s} |f_p(\boldsymbol{\alpha}; \mathbf{u}_i)|^{2s}, \end{aligned}$$

and another application of (3.7) therefore yields

$$S_4(p; \boldsymbol{\eta}) \ll p^{2sd-q_j} \max_{\mathbf{a} \in [1,p]^d} S_5(\mathbf{a}, p; \boldsymbol{\eta}), \tag{3.8}$$

where

$$S_5(\mathbf{a}, p; \boldsymbol{\eta}) = \int_{\mathbb{T}^r} |G_p(\boldsymbol{\alpha}; \boldsymbol{\eta})|^2 |f_p(\boldsymbol{\alpha}; \mathbf{a})|^{2s} \, d\boldsymbol{\alpha}.$$

Next we observe that $S_5(\mathbf{a}, p; \boldsymbol{\eta})$ is the number of solutions of the system

$$\sum_{n=1}^{r} \eta_n (\Psi_{\mathbf{i}}(\mathbf{z}_n) - \Psi_{\mathbf{i}}(\mathbf{w}_n)) = \sum_{m=1}^{s} ((p\mathbf{x}_m + \mathbf{a})^{\mathbf{i}} - (p\mathbf{y}_m + \mathbf{a})^{\mathbf{i}}) \qquad (1 \le |\mathbf{i}| \le k)$$

with $-a_l/p < x_{ml}, y_{ml} \le (Q - a_l)/p$ and with $\text{Jac}(\boldsymbol{\Psi}; \mathbf{z})$ and $\text{Jac}(\boldsymbol{\Psi}; \mathbf{w})$ both having rank r_j modulo p. By Lemma 2.3, we see that this is also equal to the number of solutions of the system

$$\sum_{n=1}^{r} \eta_n (\Phi_{\mathbf{i}}(\mathbf{z}_n) - \Phi_{\mathbf{i}}(\mathbf{w}_n)) = p^{|\mathbf{i}|} \sum_{m=1}^{s} (\mathbf{x}_m^{\mathbf{i}} - \mathbf{y}_m^{\mathbf{i}}) \qquad (1 \le |\mathbf{i}| \le k),$$

where $\Phi_{\mathbf{i}}(\mathbf{z})$ is given by (2.6). Moreover, one sees easily by applying elementary row operations that $\mathrm{Jac}(\boldsymbol{\Psi};\mathbf{z})$ and $\mathrm{Jac}(\boldsymbol{\Phi};\mathbf{z})$ have the same rank. Let us write $\boldsymbol{\alpha}\mathbf{p}$ for the r-dimensional vector whose component indexed by $\mathbf{i}$ is $\alpha_{\mathbf{i}}p^{|\mathbf{i}|}$, and put

$$t_{\mathbf{i}}(\mathbf{z};\boldsymbol{\eta}) = \eta_1\Phi_{\mathbf{i}}(\mathbf{z}_1) + \cdots + \eta_r\Phi_{\mathbf{i}}(\mathbf{z}_r).$$

Then we have

$$S_5(\mathbf{a},p;\boldsymbol{\eta}) \ll \int_{\mathbb{T}^r} |H_p(\boldsymbol{\alpha})^2 f(\boldsymbol{\alpha}\mathbf{p};QP^{-\theta})^{2s}|\,d\boldsymbol{\alpha}, \tag{3.9}$$

where

$$H_p(\boldsymbol{\alpha};\boldsymbol{\eta}) = \sum_{\mathbf{z}} e\left(\sum_{1\le|\mathbf{i}|\le k} \alpha_{\mathbf{i}} t_{\mathbf{i}}(\mathbf{z};\boldsymbol{\eta})\right),$$

and where the sum is over all $\mathbf{z}_1,\dots,\mathbf{z}_r \in [1,P]^d$ for which $\mathrm{Jac}(\boldsymbol{\Phi};\mathbf{z})$ has rank r_j modulo p. Now let $\mathcal{B}_p^*(\mathbf{u};\boldsymbol{\Phi};\boldsymbol{\eta})$ denote the set of solutions $\mathbf{z}$ modulo p^k to the system of congruences

$$t_{\mathbf{i}}(\mathbf{z};\boldsymbol{\eta}) \equiv u_{\mathbf{i}} \pmod{p^{|\mathbf{i}|}} \qquad (\mathbf{i} \succ \mathbf{j}_1)$$

with $\mathrm{Jac}(\boldsymbol{\Phi};\mathbf{z})$ of rank r_j modulo p. Put

$$H_p(\boldsymbol{\alpha};\mathbf{z};\boldsymbol{\eta}) = \sum_{\substack{\mathbf{x}\in[1,P]^{rd} \\ \mathbf{x}\equiv\mathbf{z} \ (\mathrm{mod}\ p^k)}} e\left(\sum_{1\le|\mathbf{i}|\le k} \alpha_{\mathbf{i}} t_{\mathbf{i}}(\mathbf{x};\boldsymbol{\eta})\right)$$

and

$$I_p(\boldsymbol{\alpha};\boldsymbol{\eta}) = \sum_{\mathbf{u}} \left| \sum_{\mathbf{z}\in\mathcal{B}_p^*(\mathbf{u};\boldsymbol{\Phi};\boldsymbol{\eta})} H_p(\boldsymbol{\alpha};\mathbf{z};\boldsymbol{\eta}) \right|^2,$$

where the summation is over $\mathbf{u}$ with $1 \le u_{\mathbf{i}} \le p^{|\mathbf{i}|}$ for each $\mathbf{i} \succ \mathbf{j}_1$. By Cauchy's inequality, we have

$$I_p(\boldsymbol{\alpha};\boldsymbol{\eta}) \le \sum_{\mathbf{u}} \mathrm{card}\ \mathcal{B}_p^*(\mathbf{u};\boldsymbol{\Phi};\boldsymbol{\eta}) \sum_{\mathbf{z}\in\mathcal{B}_p^*(\mathbf{u};\boldsymbol{\Phi};\boldsymbol{\eta})} |H_p(\boldsymbol{\alpha};\mathbf{z};\boldsymbol{\eta})|^2,$$

and Lemma 2.2 tells us that

$$\mathrm{card}\ \mathcal{B}_p^*(\mathbf{u};\boldsymbol{\Phi};\boldsymbol{\eta}) \ll p^{krd-K_j}.$$

Thus from (3.8) and (3.9) we finally obtain

$$\begin{aligned} S_4(p;\boldsymbol{\eta}) &\ll p^{2sd-q_j}\int_{\mathbb{T}^r} I_p(\boldsymbol{\alpha};\boldsymbol{\eta})|f(\boldsymbol{\alpha}\mathbf{p};QP^{-\theta})|^{2s}d\boldsymbol{\alpha} \\ &\ll p^{2sd+\omega(k,j,d)} \sum_{\mathbf{z}\in[1,p^k]^{rd}} \int_{\mathbb{T}^r} |H_p(\boldsymbol{\alpha};\mathbf{z};\boldsymbol{\eta})^2 f(\boldsymbol{\alpha}\mathbf{p};QP^{-\theta})^{2s}|\,d\boldsymbol{\alpha}, \end{aligned}$$

and the lemma now follows from (3.4) and (3.5) on noting that $|H_p(\boldsymbol{\alpha};\mathbf{z};\boldsymbol{\eta})| = |H_p(\boldsymbol{\alpha};\mathbf{z};\mathbf{1})|$ and considering the underlying diophantine equations. □

We now develop a differencing lemma that allows us to repeat the procedure embedded in the above result.

LEMMA 3.2. *Suppose that* $P^\theta \le Q \le P$, *write* $H = P^{1-k\theta}$, *and let* $(\boldsymbol{\Phi})$ *be a* *system of type* (j, P, A). *Then there exist* $\mathbf{h} \in [-H, H]^d \cap (\mathbb{Z} \setminus \{0\})^d$ *and* $p \in \mathcal{P}(\theta)$ *such that*

$$L_s(P, Q, \theta; \mathbf{\Phi}) \ll P^{(2d-1-(d-1)k\theta)r} J_s(QP^{-\theta}) + H^{dr} \left(K_s(P, QP^{-\theta}; \mathbf{\Upsilon}) J_s(QP^{-\theta}) \right)^{1/2},$$

where we have

$$\Upsilon_{\mathbf{i}}(\mathbf{z}) = p^{-|\mathbf{i}|}(\Phi_{\mathbf{i}}(\mathbf{z} + \mathbf{h}p^k) - \Phi_{\mathbf{i}}(\mathbf{z})) \qquad (1 \leq |\mathbf{i}| \leq k).$$

Proof. Fix a prime $p \in \mathcal{P}(\theta)$. We have $L_s(P, Q, \theta, p; \mathbf{\Phi}) = U_0 + U_1$, where U_0 denotes the number of solutions of (3.2), with $\mathbf{\Psi}$ replaced by $\mathbf{\Phi}$, for which $z_{nl} = w_{nl}$ for some n and l, and where U_1 is the number of solutions with $z_{nl} \neq w_{nl}$ for all n and l.

First of all, suppose that $U_0 \geq U_1$. In view of the congruence conditions on $\mathbf{z}$ and $\mathbf{w}$, we have

$$U_0 \ll P^{2d-1-(d-1)k\theta} \int_{\mathbb{T}^r} g_p(\boldsymbol{\alpha})^{r-1} |f(\boldsymbol{\alpha}\mathbf{p}; QP^{-\theta})|^{2s} \, d\boldsymbol{\alpha},$$

where

$$g_p(\boldsymbol{\alpha}) = \sum_{\mathbf{z} \in [1, p^k]^d} \left| \sum_{\substack{\mathbf{x} \in [1, P]^d \\ \mathbf{x} \equiv \mathbf{z} \pmod{p^k}}} e\left(\sum_{1 \leq |\mathbf{i}| \leq k} \alpha_{\mathbf{i}} \Phi_{\mathbf{i}}(\mathbf{x}) \right) \right|^2.$$

It now follows from Hölder's inequality that U_0 is bounded above by

$$P^{2d-1-(d-1)k\theta} \left(\int_{\mathbb{T}^r} g_p(\boldsymbol{\alpha})^r |f(\boldsymbol{\alpha}\mathbf{p}; QP^{-\theta})|^{2s} d\boldsymbol{\alpha} \right)^{1-1/r} \left(\int_{\mathbb{T}^r} |f(\boldsymbol{\alpha}\mathbf{p}; QP^{-\theta})|^{2s} d\boldsymbol{\alpha} \right)^{1/r},$$

so on considering the underlying diophantine equations we see that

$$L_s(P, Q, \theta, p; \mathbf{\Phi}) \ll P^{(2d-1-(d-1)k\theta)r} J_s(QP^{-\theta}). \tag{3.10}$$

Now suppose instead that $U_1 \geq U_0$. Then we can write

$$w_{nl} = z_{nl} + h_{nl} p^k \qquad (1 \leq n \leq r, \ 1 \leq l \leq d),$$

where the h_{nl} are integers satisfying $1 \leq |h_{nl}| \leq H$. We therefore see that U_1 is bounded above by the number of solutions of the system

$$\sum_{n=1}^{r} \Upsilon_{\mathbf{i}}(\mathbf{z}_n; \mathbf{h}_n; p) = \sum_{m=1}^{s} (\mathbf{u}_m^{\mathbf{i}} - \mathbf{v}_m^{\mathbf{i}}) \qquad (1 \leq |\mathbf{i}| \leq k)$$

with $\mathbf{z}_n \in [1, P]^d$, with $\mathbf{h}_n$ as above, and with $\mathbf{u}_m, \mathbf{v}_m \in [1, QP^{-\theta}]^d$. Now write

$$W_p(\boldsymbol{\alpha}; \mathbf{h}) = \sum_{\mathbf{z} \in [1, P]^d} e\left(\sum_{1 \leq |\mathbf{i}| \leq k} \alpha_{\mathbf{i}} \Upsilon_{\mathbf{i}}(\mathbf{z}; \mathbf{h}; p) \right). \tag{3.11}$$

Then we have

$$U_1 \leq \int_{\mathbb{T}^r} \left(\sum_{\mathbf{h}} W_p(\boldsymbol{\alpha}; \mathbf{h}) \right)^r |f(\boldsymbol{\alpha}; QP^{-\theta})|^{2s} \, d\boldsymbol{\alpha},$$

where the summation is over $h_1, \ldots, h_d$ with $1 \leq |h_l| \leq H$. Furthermore, by Hölder's inequality, one has

$$\left(\sum_{\mathbf{h}} W_p(\boldsymbol{\alpha};\mathbf{h})\right)^r \ll H^{d(r-1)} \sum_{\mathbf{h}} |W_p(\boldsymbol{\alpha};\mathbf{h})|^r.$$

Thus, by applying the Cauchy-Schwarz inequalities, we deduce that U_1 is bounded above by

$$\left(H^{2d(r-1)+d}\int_{\mathbb{T}^r} \sum_{\mathbf{h}} |W_p(\boldsymbol{\alpha};\mathbf{h})^{2r} f(\boldsymbol{\alpha};QP^{-\theta})^{2s}|\,d\boldsymbol{\alpha}\right)^{1/2}\left(\int_{\mathbb{T}^r} |f(\boldsymbol{\alpha};QP^{-\theta})|^{2s}\,d\boldsymbol{\alpha}\right)^{1/2},$$

and the integral in the first factor is bounded by $H^d K_s(P, QP^{-\theta};\boldsymbol{\Upsilon})$, where $\Upsilon_{\mathbf{i}} = \Upsilon_{\mathbf{i}}(\mathbf{z};\mathbf{h};p)$ for some $\mathbf{h}$. The lemma now follows on recalling (3.10) and taking the maximum over primes $p \in \mathcal{P}(\theta)$. □

In order to use Lemmas 3.1 and 3.2, we must describe the polynomials $\Psi_{\mathbf{i}}$ to which we want to apply these results and then verify that they satisfy the conditions of the lemmas. To this end, we first define the difference operator Δ_j recursively by

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

and

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

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

$$\Psi_{\mathbf{i},j+1}(\mathbf{z};\mathbf{h};\mathbf{p}) = p_{j+1}^{-|\mathbf{i}|}\Delta_1(\Phi_{\mathbf{i}}(\mathbf{z};\boldsymbol{\Psi}_j(\mathbf{z};\mathbf{h}_1,\ldots,\mathbf{h}_j;p_1,\ldots,p_j));\mathbf{h}_{j+1}p_{j+1}^k),$$

where the polynomials $\Phi_{\mathbf{i}}(\mathbf{z};\boldsymbol{\Psi})$ are defined by (2.6) and where we have written $\boldsymbol{\Psi}_j$ for the set of all $\Psi_{\mathbf{i},j}$ with $1 \le |\mathbf{i}| \le k$. Since the terms of highest degree in $\Psi_{\mathbf{i},j}(\mathbf{z})$ and $\Phi_{\mathbf{i}}(\mathbf{z};\boldsymbol{\Psi}_j)$ are identical, we have

$$\Psi_{\mathbf{i},j}(\mathbf{z};\mathbf{h};\mathbf{p}) = (p_1\cdots p_j)^{-|\mathbf{i}|}\Delta_j(\mathbf{z}^{\mathbf{i}};\mathbf{h}_1p_1^k,\ldots,\mathbf{h}_jp_j^k) + E(\mathbf{z};\mathbf{h};\mathbf{p}), \tag{3.12}$$

where $E(\mathbf{z};\mathbf{h};\mathbf{p})$ has total degree strictly less than $|\mathbf{i}|-j$ in the variables $z_1,\ldots,z_d$. We typically think of $\mathbf{h}$ and $\mathbf{p}$ as fixed and regard $\Psi_{\mathbf{i},j}$ as a polynomial in $\mathbf{z}$. When $\mathbf{h} = (\mathbf{h}_1,\ldots,\mathbf{h}_j)$ 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 $\mathbf{h}_l^* = (h_{1l},\ldots,h_{jl})$. We start by relating our vector difference operator to the more familiar scalar one. When $\mathcal{A} = \{i_1,\ldots,i_m\}$ and $\mathcal{B} = \{j_1,\ldots,j_t\}$ with $\mathcal{A}\cap\mathcal{B} = \emptyset$, we write

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

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

Lemma 3.3. *One has*

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

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

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

It follows that

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

and this gives the result. □

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 $p_1, \dots, p_j \in \mathbb{Z}$ have the property that $0 < |h_{nl} p_n^k| \le cP$ 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 [**9**], 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 (3.12) and Lemma 3.3 that the terms of highest degree in $\Psi_{\mathbf{i},j}(\mathbf{z};\mathbf{h};\mathbf{p})$, which we denote by $G_{\mathbf{i},j}(\mathbf{z})$, are given by

$$(p_1\cdots p_j)^{-|\mathbf{i}|}\sum_{\mathcal{A}_1\sqcup\cdots\sqcup\mathcal{A}_d=\{1,\ldots,j\}}\left(\prod_{l=1}^d \frac{i_l!}{(i_l-|\mathcal{A}_l|)!}\prod_{n\in\mathcal{A}_l} h_{nl}p_n^k\right) z_1^{i_1-|\mathcal{A}_1|}\cdots z_d^{i_d-|\mathcal{A}_d|}.$$

Conditions (1), (2), and (3) of Definition 2.4 follow immediately. To check condition (4), we fix $\mathbf{i}$ with $\mathbf{i}\succ\mathbf{j}_1$ (so in particular $i_1\ge j$) and consider the term $z_1^{i_1-j}z_2^{i_2}\cdots z_d^{i_d}$ arising from the choice $\mathcal{A}_1=\{1,\ldots,j\}$ in the expression for $G_{\mathbf{i},j}(\mathbf{z})$ above. Suppose now that there is some $\mathbf{i}'$ with $|\mathbf{i}'|=|\mathbf{i}|$ 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 $\mathcal{A}_1=\{1,\ldots,j\}$, and it follows that $\mathbf{i}'=\mathbf{i}$. Otherwise, we must have $i_1'<i_1$, which implies that $\mathbf{i}'\prec\mathbf{i}$. □

Note that in our applications we can take $c=2^k$ in the above lemma, since the prime p used at each stage satisfies $P^\theta<p\le 2P^\theta$ for some $\theta\le 1/k$, and the corresponding values of $h_1,\ldots,h_d$ are bounded in modulus by $H=P^{1-k\theta}$. Starting with $j=0$, we apply Lemma 3.1 to the system $(\mathbf{\Psi})=(\mathbf{\Psi}_j)$ and then apply Lemma 3.2 with $(\mathbf{\Phi})=(\mathbf{\Phi}_j)$, where $\mathbf{\Phi}_j=\mathbf{\Phi}(\mathbf{\Psi}_j)$ is given by (2.6). This puts us in position to apply Lemma 3.1 again with $(\mathbf{\Psi})$ replaced by the system $(\mathbf{\Upsilon})=(\mathbf{\Psi}_{j+1})$ and hence to repeat the process.

4. *Mean value theorems*

By using only first differences, one obtains the following simple result, which is useful for generating some preliminary admissible exponents. When Δ_s is an admissible exponent, we sometimes refer to the quantity $\lambda_s=2sd-K+\Delta_s$ as a permissible exponent.

Theorem 4.1. *If Δ_s is an admissible exponent satisfying $\Delta_s\le(k-1)(r+1)$, then the exponent $\Delta_{s+r}=\Delta_s(1-1/k)$ is also admissible.*

Proof. By Lemma 3.1, we have

$$K_s(P,P;\mathbf{\Psi}_0)\ll P^{2rd-(1-\theta)(r+1)}J_s(P)+P^{\theta(2sd+krd-K)}L_s(P,P,\theta;\mathbf{\Phi}_0),$$

and the argument of the proof of Lemma 3.2 gives

$$L_s(P,P,\theta;\mathbf{\Phi}_0)\ll P^{(2d-1-(d-1)k\theta)r}J_s(P^{1-\theta})+\int_{\mathbb{T}^r}\left|\sum_{\mathbf{h}}W_p(\boldsymbol{\alpha};\mathbf{h})\right|^r|f(\boldsymbol{\alpha};P^{1-\theta})|^{2s}\,d\boldsymbol{\alpha}$$

for some $p\in\mathcal{P}(\theta)$, where $W_p(\boldsymbol{\alpha};\mathbf{h})$ is as in (3.11) and where the summation is over $\mathbf{h}\in[-H,H]^d$. Taking $\theta=1/k$ gives $H=1$, so after making a trivial estimate we find that

$$L_s(P,P,\theta;\mathbf{\Phi}_0)\ll P^{rd}J_s(P^{1-\theta}),$$

whence

$$K_s(P,P;\mathbf{\Psi}_0)\ll P^{2rd-(1-\theta)(r+1)}J_s(P)+P^{rd+\theta(2sd+krd-K)}J_s(P^{1-\theta}).$$

Suppose that the exponent $\lambda_s=2sd-K+\Delta_s$ is permissible, where one has $\Delta_s\le(k-1)(r+1)$. Then we have

$$J_{s+r}(P)\ll K_s(P,P;\mathbf{\Psi}_0)\ll P^{\Lambda_1}+P^{\Lambda_2},$$

where

$$\Lambda_1 = 2(s+r)d - K + \Delta_s - (1-\theta)(r+1)$$

and

$$\Lambda_2 = rd + \theta(2sd + krd - K) + (1-\theta)\lambda_s = 2(s+r)d - K + \Delta_s(1-\theta).$$

The inequality $\Delta_s \le (k-1)(r+1)$ shows that $\Lambda_1 \le \Lambda_2$, and the exponent $\Delta_{s+r} = (1-1/k)\Delta_s$ is therefore admissible. □

We can obtain somewhat stronger results via repeated differencing. The following theorem, while not in a form convenient for direct application, provides our sharpest admissible exponents for large values of s and k. In stating our theorem, we shall find it convenient to introduce the notation

$$\Omega_J = K - K_J - q_J, \tag{4.1}$$

which we loosely view as a measurement of the loss of potential congruence information suffered at the Jth difference.

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

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

and

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

Finally, we set

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

and

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

Then Δ_s is an admissible exponent for $s = u + lr$ for all positive integers l.

Proof. Let us initially fix $s \ge u + r$, and suppose that λ_s is a permissible exponent. In view of the hypothesis on Δ_u, we may clearly suppose that $\Delta_s = \lambda_s - 2sd + K \le (k-1)(r+1)$. Take j to be the least integer for which $\phi(j, s+r, 1) = \theta_{s+r}$, and write $\phi_J = \phi(j, s+r, J)$ for $J = j,\dots,1$. Also note that the minimality of j ensures that $\phi_J < 1/k$ whenever $J < j$. We adopt the notation

$$M_i = P^{\phi_i}, \qquad H_i = PM_i^{-k}, \qquad Q_i = P(M_1\cdots M_i)^{-1} \qquad (1 \le i \le j),$$

with the convention that $Q_0 = P$. We first show inductively that

$$L_s(P, Q_J, \phi_{J+1}; \mathbf{\Phi}_J) \ll P^{(2d-1-(d-1)k\phi_{J+1})r} Q_{J+1}^{\lambda_s} \tag{4.4}$$

for each $J = j-1,\dots,0$. First of all, Lemma 3.2 gives

$$L_s(P, Q_{j-1}, \phi_j; \mathbf{\Phi}_{j-1}) \ll P^{(2d-1-(d-1)k\phi_j)r} J_s(Q_j) + H_j^{dr}(K_s(P, Q_j; \mathbf{\Psi}_j) J_s(Q_j))^{1/2}.$$

Since $\phi_j = 1/k$, we have $H_j = 1$, so a trivial estimate yields

$$L_s(P, Q_{j-1}, \phi_j; \mathbf{\Phi}_{j-1}) \ll P^{dr} Q_j^{\lambda_s},$$

and (4.4) follows in the case $J = j-1$. Now suppose that (4.4) holds for J. Then by Lemmas 3.1 and 3.2 we have

$$L_s(P, Q_{J-1}, \phi_J; \mathbf{\Phi}_{J-1}) \ll P^{(2d-1-(d-1)k\phi_J)r} J_s(Q_J) + H_J^{dr} Q_J^{\lambda_s} (T_1 + T_2)^{1/2},$$

where

$$T_1 = P^{2rd-r-1} M_{J+1}^{r+1} \quad \text{and} \quad T_2 = P^{2rd-r} M_{J+1}^{2sd+\omega(k,J,d)-r(d-1)k-\lambda_s}.$$

A simple calculation reveals that $T_1 \le T_2$, provided that an exponent Δ_s satisfying $\Delta_s \le (k-1)(r+1) + \Omega_J$ is admissible, and this latter inequality follows from our earlier remarks, since it is clear from (4.1) that $\Omega_J \ge 0$. Thus we find that

$$L_s(P, Q_{J-1}, \phi_J; \mathbf{\Phi}_{J-1}) \ll Q_J^{\lambda_s} (P^{\Lambda_1} + P^{\Lambda_2}),$$

where

$$\Lambda_1 = (2d - 1 - (d-1)k\phi_J)r$$

and

$$\Lambda_2 = dr(1 - k\phi_J) + dr - \frac{r}{2} + \frac{\phi_{J+1}}{2}(2sd + \omega(k,J,d) - r(d-1)k - \lambda_s).$$

It follows with a little computation from (4.1), (4.2), and our initial remarks that in fact $\Lambda_1 = \Lambda_2$, and we therefore obtain (4.4) with J replaced by $J-1$. We now apply (4.4) with $J = 0$ to conclude that

$$L_s(P, P, \phi_1; \mathbf{\Phi}_0) \ll P^{(2d-1-(d-1)k\phi_1)r + (1-\phi_1)\lambda_s}.$$

Thus Lemma 3.1 gives

$$J_{s+r}(P) \ll K_s(P, P; \mathbf{\Psi}_0) \ll P^{\Lambda_3} + P^{\Lambda_4},$$

where

$$\Lambda_3 = 2rd - (1 - \phi_1)(r+1) + \lambda_s$$

and

$$\Lambda_4 = \phi_1(2sd + krd - K) + (2d - 1 - (d-1)k\phi_1)r + (1-\phi_1)\lambda_s.$$

Since $\Delta_s \le (k-1)(r+1)$, we find after a short computation that $\Lambda_3 \le \Lambda_4$, whence the exponent

$$\begin{aligned} \lambda_{s+r} &= \phi_1(2sd + krd - K) + (2d - 1 - (d-1)k\phi_1)r + (1-\phi_1)\lambda_s \\ &= 2(s+r)d - K + \Delta_s(1-\phi_1) + r(k\phi_1 - 1) \end{aligned}$$

is permissible. The theorem follows by induction on recalling that $\phi_1 = \theta_{s+r}$. $\square$

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**], [**11**], [**12**], 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 $s \ge 2r$ and that Δ_{s-r} is an admissible exponent* *satisfying* $r(\log k)^2 < \Delta_{s-r} \le (k-1)(r+1)$. *Write* $\delta_{s-r} = \Delta_{s-r}/(rk)$, *and define* δ_s *to be the unique (positive) solution of the equation*

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

Then the exponent $\Delta_s = rk\delta_s$ *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

$$\Delta_s = \Delta_{s-r}(1-\theta_s) + r(k\theta_s - 1) = rk\delta_{s-r} - r + rk\theta_s(1-\delta_{s-r}) \tag{4.6}$$

is admissible. For $0 \le J < j$, we see from (4.1) and Lemma 2.1 that

$$\Omega_J \le \frac{dk}{d+1}\left[\binom{k+d}{d} - \binom{k-j+d}{d}\right] \le r(\log k)^{1/2}$$

for k sufficiently large. Thus on writing ϕ_J for $\phi(j,s,J)$, we deduce from (4.2) that

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

where

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

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

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

so in particular we have

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

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

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

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

$$\frac{\Delta_s}{rk} \le \delta_{s-r}\left(1 - \frac{2-w}{k(1+\delta_{s-r})}\right),$$

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

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

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), it must be the case that $\Delta_s/(rk) \le \delta_s$, and it follows that $rk\delta_s$ is an admissible exponent. □

LEMMA 4.4. *If $k > d+1$, then the exponent $\Delta_{4r} = r(k-2)$ is admissible.*

Proof. First of all, the exponent $\Delta_r = K$ is trivially admissible, and it follows easily from Lemma 2.1 that $K \le (k-1)(r+1)$ whenever $k \ge d+1$. Thus we may apply Theorem 4.1 successively to deduce that the exponent $\Delta_{4r} = K(1-\frac{1}{k})^3$ is admissible, and one has

$$\Delta_{4r} \le rk(1-\tfrac{1}{k})^3 \le rk - 3r(1-\tfrac{1}{k}) \le r(k-2)$$

whenever $k \ge 3$, which establishes the lemma. □

On combining Lemma 4.4 with Theorem 4.1, we can produce admissible exponents Δ_s satisfying $\Delta_s \ll rke^{-s/rk}$. We are now in a position to state the stronger mean value estimates arising from repeated differencing in a form convenient for application.

THEOREM 4.5. *Suppose that k is sufficiently large in terms of d, define s_0 and s_1 as in* (1.12) *and* (1.13)*, and write $L = (\log k)^2$. Then the exponents Δ_s defined by*

$$\Delta_s = \begin{cases} rke^{2-2s/rk} & \text{if } 1 \le s \le s_0, \\ e^{2+2/k}rL\left(1-\frac{3}{2k}(1-\frac{d}{2L})\right)^{(s-s_0)/r} & \text{if } s_0 < s \le s_1 \end{cases}$$

are admissible.

Proof. We define δ_s to be the unique positive solution of the equation

$$\delta_s + \log\delta_s = 1 - \frac{2(s-4r)}{rk} + \frac{2(s-4r)}{rk(\log k)^{3/2}}. \tag{4.10}$$

We show inductively that the exponent $\Delta_s = rk\delta_s$ is admissible whenever $4r < s \le s_0$. First of all, suppose that $4r < s \le 5r$. Then by Lemma 4.4 we know that $\Delta_s^* = r(k-2)$ is admissible, and furthermore

$$\frac{\Delta_s^*}{rk} + \log\frac{\Delta_s^*}{rk} < 1 - \frac{2}{k} < \delta_s + \log\delta_s,$$

since $0 < s - 4r \le r$. It follows that $\Delta_s^*/(rk) < \delta_s$, and hence $\Delta_s = rk\delta_s$ is admissible. Now suppose that $5r < s \le s_0$ and that the exponent $\Delta_{s-r} = rk\delta_{s-r}$ is admissible. Then we have

$$\delta_{s-r} + \log\delta_{s-r} > 1 - \frac{2(s_0-4r)}{rk} > 1 - \log k + 2\log\log k.$$

Since $\delta_{s-r} < 1$, we deduce that $\delta_{s-r} > (\log k)^2/k$, and thus

$$\Delta'_{s-r} = \min\{\Delta_{s-r},\, (k-1)(r+1)\}$$

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

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

and where $\delta'_{s-r} = \Delta'_{s-r}/(rk) \le \delta_{s-r}$. On applying (4.10) with s replaced by $s-r$, we find that $\gamma_s + \log\gamma_s \le \delta_s + \log\delta_s$, and hence $\gamma_s \le \delta_s$. Thus $\Delta_s = rk\delta_s$ is admissible.

To complete the proof of the theorem, we first note that the result holds trivially for $1 \le s \le 4r$, since $K \le rk$. For $4r < s \le s_0$, we see from (4.10) that

$$\log \delta_s \le 2 - \frac{2s}{rk},$$

provided that k is sufficiently large. Finally, if $s > s_0$, we take t to be the integer with $s_0 - r < t \le s_0$ and $t \equiv s \pmod r$. Then we know that $\Delta_t = rke^{2-2t/rk}$ is an admissible exponent, and we have

$$e^2 r(\log k)^2 \le \Delta_t < e^{2+2/k} r(\log k)^2. \tag{4.11}$$

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

$$\phi^*(2, t + r, 1) = \frac{1}{2k} + \left(\frac{1}{2} + \frac{\Omega_1 - \Delta_t}{2kr}\right)\frac{1}{k} = \frac{1}{k} + \frac{\Omega_1 - \Delta_t}{2k^2 r}.$$

It therefore follows from (4.3) that the exponent

$$\Delta_{t+r} = \Delta_t \left(1 - \frac{3}{2k} + \frac{\Delta_t - \Omega_1}{2k^2 r}\right) + \frac{\Omega_1}{2k} \tag{4.12}$$

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

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

is admissible, and the theorem follows on substituting (4.11) and recalling that $t \le s_0$. □

To deduce Theorem 1.1, we note that

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

for $k \ge 6$, and thus

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

Theorem 1.1 now follows immediately from Theorem 4.5 when $s \le s_1$. For $s > s_1$, it follows from Theorem 1.3 that we may take $\Delta_s = 0$, so Theorem 1.1 holds in that case as well (but is of little value). The remainder of the paper is largely devoted to the proof of Theorem 1.3, which uses Theorem 1.1 only with $s \le s_1$.

5. *Weyl-type estimates*

In this section, we aim to deduce Theorem 1.2 from the bounds provided by Theorem 1.1. Our strategy combines ideas of Vaughan and Baker and actually leads to a result (Theorem 5.5) containing somewhat more information than given by Theorem 1.2. We first use the large sieve as in Vaughan [**9**] to get a preliminary estimate in terms of a rational approximation to some $\alpha_{\mathbf{j}}$. We then use this result within a similar argument devised by Baker to control the least common multiple of the denominators of the various rational approximations. This strategy gives information only about the $\alpha_{\mathbf{j}}$ with $|\mathbf{j}| \geq 2$, so the remaining ingredient is an analogue of Baker's final coefficient lemma. Our preliminary estimate is as follows.

THEOREM 5.1. *Fix* $\mathbf{j}$ *with* $2 \leq |\mathbf{j}| \leq k$, *and let* $q \geq 1$ *and* a *be relatively prime integers satisfying* $|q\alpha_{\mathbf{j}} - a| \leq q^{-1}$. *Further, let* s *be any positive integer, and let* $\Delta = \Delta_{s,k-1,d}$ *denote an admissible exponent for* $(s, k-1, d)$. *Then one has*

$$|f(\boldsymbol{\alpha})| \ll P^d [P^{\Delta}(qP^{-|\mathbf{j}|} + P^{-1} + q^{-1})]^{1/2s} \log P.$$

Proof. We follow the argument of the proof of Vaughan [**9**], Theorem 5.2. We start by performing a Weyl shift with respect to the variable x_1. Consider a set $\mathcal{M} \subseteq [1, P] \cap \mathbb{Z}$ with $|\mathcal{M}| = M$. Then, for any $m \in \mathcal{M}$, one has

$$\begin{aligned} f(\boldsymbol{\alpha}) &= \sum_{\substack{x_1 \in [1+m, P+m] \\ x_2, \ldots, x_d \in [1,P]}} e\left(\sum_{1 \leq |\mathbf{i}| \leq k} \alpha_{\mathbf{i}} (x_1 - m)^{i_1} x_2^{i_2} \cdots x_d^{i_d} \right) \\ &= \int_0^1 \sum_{\substack{x_1 \in [1, 2P] \\ x_2, \ldots, x_d \in [1,P]}} e\left(\sum_{1 \leq |\mathbf{i}| \leq k} \alpha_{\mathbf{i}} (x_1 - m)^{i_1} x_2^{i_2} \cdots x_d^{i_d} + x_1 \beta \right) \sum_{y=1+m}^{P+m} e(-y\beta)\, d\beta. \end{aligned}$$

Summing over all $m \in \mathcal{M}$, we find that

$$M|f(\boldsymbol{\alpha})| \ll \int_0^1 \sum_{m \in \mathcal{M}} |g(m, \beta)| \min(P, ||\beta||^{-1})\, d\beta,$$

where

$$g(m, \beta) = \sum_{\substack{x_1 \in [1, 2P] \\ x_2, \ldots, x_d \in [1,P]}} e\left(\sum_{1 \leq |\mathbf{i}| \leq k} \alpha_{\mathbf{i}} (x_1 - m)^{i_1} x_2^{i_2} \cdots x_d^{i_d} + x_1 \beta \right).$$

It follows that

$$|f(\boldsymbol{\alpha})| \ll M^{-1} \left(\sup_{\beta \in [0,1]} \sum_{m \in \mathcal{M}} |g(m, \beta)| \right) \log P,$$

and an application of Hölder's inequality yields

$$|f(\boldsymbol{\alpha})|^{2s} \ll M^{-1} (\log P)^{2s} \sum_{m \in \mathcal{M}} |g(m, \beta)|^{2s} \tag{5.1}$$

for some $\beta \in [0, 1]$. We now aim to express $|g(m, \beta)|^{2s}$ in a form to which a version of the large sieve can be applied. By the binomial theorem, we have

$$\sum_{1 \leq |\mathbf{i}| \leq k} \alpha_{\mathbf{i}} (x_1 - m)^{i_1} x_2^{i_2} \cdots x_d^{i_d} = \sum_{1 \leq |\mathbf{i}| \leq k} \alpha_{\mathbf{i}} \sum_{j_1 = 0}^{i_1} \binom{i_1}{j_1} (-m)^{i_1 - j_1} x_1^{j_1} x_2^{i_2} \cdots x_d^{i_d}.$$

We first split off the terms on the right for which $j_1 + i_2 + \cdots + i_d \in \{0, k\}$. Then on writing $\mathbf{j} = (j_1, i_2, \ldots, i_d)$ and interchanging the order of summation in the remaining terms, one finds that

$$\sum_{1 \leq |\mathbf{i}| \leq k} \alpha_{\mathbf{i}} (x_1 - m)^{i_1} x_2^{i_2} \cdots x_d^{i_d} = \sum_{|\mathbf{i}| = k} \alpha_{\mathbf{i}} \mathbf{x}^{\mathbf{i}} + \sum_{i=1}^{k} \alpha_{i,0,\ldots,0} (-m)^i + \sum_{1 \leq |\mathbf{j}| \leq k-1} \gamma_{\mathbf{j}}(m) \mathbf{x}^{\mathbf{j}},$$

where

$$\gamma_{\mathbf{j}}(m)=\sum_{i=j_1}^{k-(j_2+\cdots+j_d)}\binom{i}{j_1}\alpha_{i,j_2,\ldots,j_d}(-m)^{i-j_1}. \tag{5.2}$$

Write

$$\Gamma=\{\boldsymbol{\gamma}(m):m\in\mathcal{M}\}\subseteq\mathbb{R}^q,$$

where $q=\binom{k-1+d}{d}-1$, and

$$\mathcal{N}=\prod_{1\le|\mathbf{j}|\le k-1}\mathcal{N}_{\mathbf{j}},$$

where $\mathcal{N}_{\mathbf{j}}=[1,2^{j_1}sP^{|\mathbf{j}|}]\cap\mathbb{Z}$. We also write $N_{\mathbf{j}}=|\mathcal{N}_{\mathbf{j}}|$.

Suppose that for every $x,y\in\mathcal{M}$ with $x\ne y$ one has $||\gamma_{\mathbf{j}}(x)-\gamma_{\mathbf{j}}(y)||>\delta_{\mathbf{j}}$ for some $\mathbf{j}$ with $1\le|\mathbf{j}|\le k-1$. We then have

$$\sum_{m\in\mathcal{M}}|g(m,\beta)|^{2s}\le\sum_{\boldsymbol{\gamma}\in\Gamma}\left|\sum_{\mathbf{n}\in\mathcal{N}}a(\mathbf{n})e(\boldsymbol{\gamma}\cdot\mathbf{n})\right|^2, \tag{5.3}$$

where

$$a(\mathbf{n})=\sideset{}{'}\sum_{\mathbf{x}_1,\ldots,\mathbf{x}_s}e\left(\sum_{|\mathbf{i}|=k}\alpha_{\mathbf{i}}(\mathbf{x}_1^{\mathbf{i}}+\cdots+\mathbf{x}_s^{\mathbf{i}})+\beta(x_{11}+\cdots+x_{s1})\right),$$

and where $\sum'$ denotes the summation over $\mathbf{x}_1,\ldots,\mathbf{x}_s\in[1,2P]\times[1,P]^{d-1}$ satisfying the system

$$\mathbf{x}_1^{\mathbf{j}}+\cdots+\mathbf{x}_s^{\mathbf{j}}=n_{\mathbf{j}}\qquad(1\le|\mathbf{j}|\le k-1).$$

Notice also that one has

$$\sum_{\mathbf{n}\in\mathcal{N}}|a(\mathbf{n})|^2\le J_{s,k-1,d}(2P). \tag{5.4}$$

Then by a q-dimensional version of the large sieve inequality (see for example Vaughan [**9**], Lemma 5.3), one deduces from (5.3) and (5.4) that

$$\sum_{m\in\mathcal{M}}|g(m,\beta)|^{2s}\ll\left(\prod_{1\le|\mathbf{j}|\le k-1}(N_{\mathbf{j}}+\delta_{\mathbf{j}}^{-1})\right)J_{s,k-1,d}(2P). \tag{5.5}$$

It therefore remains to analyze the spacing of the $\gamma_{\mathbf{j}}(m)$ defined by (5.2) as m runs through a suitably chosen set $\mathcal{M}$. For this, we need to make use of rational approximations, so let us fix $\mathbf{j}$ as in the statement of the theorem with $2\le|\mathbf{j}|\le k$. Without loss of generality, we may suppose that $j_1\ge1$, and we temporarily adopt the notation $J=j_2+\cdots+j_d$. We also fix $x,y\in\mathcal{M}$ with $x\ne y$. When $0\le j\le k-1-J$, we write $\gamma_j(m)=\gamma_{j,j_2,\ldots,j_d}(m)$ and define

$$\begin{aligned}\tau_j&=(k-J)!(\gamma_j(x)-\gamma_j(y))\\&=(k-J)!\sum_{h=j}^{k-J}\binom{h}{j}((-x)^{h-j}-(-y)^{h-j})\alpha_{h,j_2,\ldots,j_d}=\sum_{h=1}^{k-1-J}\beta_h a_{hj},\end{aligned}$$

where $a_{hj}=0$ if $h<j$,

$$a_{hj}=\frac{(k-J)!}{h+1}\binom{h+1}{j}\frac{(-x)^{h+1-j}-(-y)^{h+1-j}}{y-x}\qquad(j\le h\le k-1-J),$$

and

$$\beta_h = \alpha_{h+1,j_2,\dots,j_d}(h+1)(y-x).$$

Thus, by applying the argument leading to inequality (5.33) of Vaughan [**9**], with k replaced by $k-J$, we may conclude that

$$||((k-J)!)^{k-J}\alpha_{j,j_2,\dots,j_d}(x-y)|| \ll \sum_{h=j-1}^{k-1-J} ||\gamma_h(x)-\gamma_h(y)||P^{h-j+1}$$

for all j with $1 \le j \le k-J$. In particular, on returning to our original notation, it follows that

$$||(k!)^k\alpha_{\mathbf{j}}(x-y)|| \ll ||\gamma_{\mathbf{h}}(x)-\gamma_{\mathbf{h}}(y)||P^{h_1-j_1+1} \tag{5.6}$$

for some $\mathbf{h}=(h_1,j_2,\dots,j_d)$ with $1 \le |\mathbf{j}|-1 \le h_1+j_2+\dots+j_d \le k-1$. Here our assumption that $j_1 \ge 1$ ensures that $h_1 \ge 0$. Now suppose that $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ are coprime integers with $|q\alpha_{\mathbf{j}}-a| \le q^{-1}$, and write $N=\min(P,q)$. Fix $x \in [1,N]$. If $y \in [1,N]$ satisfies

$$||(k!)^k\alpha_{\mathbf{j}}(x-y)|| \le P^{1-|\mathbf{j}|},$$

then by the triangle inequality one has

$$||(k!)^k a(x-y)/q|| \le P^{1-|\mathbf{j}|}+(k!)^kNq^{-2}.$$

Hence the number of choices for the residue class modulo q of $(k!)^ky$ is at most $2qP^{1-|\mathbf{j}|}+2(k!)^kNq^{-1}+1$, so the number of possibilities for $y \in [1,N]$ is at most

$$R=((k!)^kNq^{-1}+1)(2qP^{1-|\mathbf{j}|}+2(k!)^kNq^{-1}+1).$$

It follows that there exists a set $\mathcal{M} \subseteq [1,N]\cap\mathbb{Z}$ with $|\mathcal{M}|=M \ge N/(R+1)$ with the property that for any $x,y \in \mathcal{M}$ with $x \ne y$ one has $||(k!)^k\alpha_{\mathbf{j}}(x-y)|| > P^{1-|\mathbf{j}|}$. In this case, (5.6) implies that $||\gamma_{\mathbf{h}}(x)-\gamma_{\mathbf{h}}(y)|| \gg P^{-|\mathbf{h}|}$ for some $\mathbf{h}$ with $1 \le |\mathbf{h}| \le k-1$. Thus we can take $\delta_{\mathbf{j}}=P^{-|\mathbf{j}|}$ in (5.5) and combine this with (5.1) to obtain

$$|f(\boldsymbol{\alpha})|^{2s} \ll M^{-1}(\log P)^{2s}P^{K-L}J_{s,k-1,d}(2P) \ll P^{2sd+\Delta}M^{-1}(\log P)^{2s},$$

where K and L are as in (1.10) and (1.14). Finally, we note that

$$M^{-1} \ll R/N \ll N^{-1}(qP^{1-|\mathbf{j}|}+Nq^{-1}+1) \ll qP^{-|\mathbf{j}|}+P^{-1}+q^{-1},$$

from which the theorem now follows. □

We now use Theorem 5.1 to show that, when $|f(\boldsymbol{\alpha})|$ is large, each coefficient $\alpha_{\mathbf{j}}$ with $2 \le |\mathbf{j}| \le k$ has a good rational approximation such that the least common multiple of the denominators is relatively small. The following theorem is modeled on Theorem 4.3 of Baker [**2**].

Theorem 5.2. *Let s be a positive integer, and let $\Delta=\Delta_{s,k-1,d}$ be an admissible exponent for $(s,k-1,d)$. Further suppose that $|f(\boldsymbol{\alpha})| \gg A$, where $Q=P^{\Delta}(P^dA^{-1})^{2s}$ satisfies $Q \ll P^{1-2\varepsilon}$ for some $\varepsilon>0$. Then there exist integers $a_{\mathbf{j}}$ and natural numbers $q_{\mathbf{j}}$, with $(q_{\mathbf{j}},a_{\mathbf{j}})=1$, such that*

$$|q_{\mathbf{j}}\alpha_{\mathbf{j}}-a_{\mathbf{j}}| \le QP^{-|\mathbf{j}|+\varepsilon} \qquad (2 \le |\mathbf{j}| \le k).$$

Moreover, the least common multiple q_0 of the numbers $q_{\mathbf{j}}$ satisfies $q_0 \ll Q(\log P)^{2s}$.

Proof. For each $\mathbf{j}$ with $2 \le |\mathbf{j}| \le k$, we may apply Dirichlet's Theorem to obtain coprime integers $q_{\mathbf{j}}$ and $a_{\mathbf{j}}$ with

$$1 \le q_{\mathbf{j}} \le Q^{-1}P^{|\mathbf{j}|-\varepsilon} \quad \text{and} \quad |q_{\mathbf{j}}\alpha_{\mathbf{j}} - a_{\mathbf{j}}| \le QP^{-|\mathbf{j}|+\varepsilon}. \tag{5.7}$$

Then by Theorem 5.1, we have

$$A^{2s} \ll |f(\boldsymbol{\alpha})|^{2s} \ll P^{2sd+\Delta}(q_{\mathbf{j}}P^{-|\mathbf{j}|} + P^{-1} + q_{\mathbf{j}}^{-1})(\log P)^{2s}$$

for each such $\mathbf{j}$. Thus we have

$$Q^{-1}(\log P)^{-2s} \ll q_{\mathbf{j}}P^{-|\mathbf{j}|} + P^{-1} + q_{\mathbf{j}}^{-1} \ll Q^{-1}P^{-\varepsilon} + q_{\mathbf{j}}^{-1},$$

and it follows that

$$q_{\mathbf{j}} \ll Q(\log P)^{2s} \ll P^{1-\varepsilon}. \tag{5.8}$$

Now fix an integer $x \in [1, P]$, and suppose there is an integer $y \in [1, P]$ such that

$$||(k!)^k\alpha_{\mathbf{j}}(x-y)|| \le P^{1-|\mathbf{j}|} \qquad (2 \le |\mathbf{j}| \le k).$$

Then by (5.7), (5.8), and the triangle inequality, one has

$$||(k!)^k a_{\mathbf{j}}(x-y)/q_{\mathbf{j}}|| \le P^{1-|\mathbf{j}|} + (k!)^k q_{\mathbf{j}}^{-1}QP^{1-|\mathbf{j}|+\varepsilon} < q_{\mathbf{j}}^{-1},$$

and it follows that $q_{\mathbf{j}}$ divides $(k!)^k a_{\mathbf{j}}(x-y)$ for each $\mathbf{j}$. Since $(q_{\mathbf{j}}, a_{\mathbf{j}}) = 1$, we deduce that q_0 divides $(k!)^k(x-y)$, and hence there are at most $R = (k!)^k P q_0^{-1} + 1$ possible choices for y. Thus there is a set of integers $\mathcal{M} \subseteq [1, P]$ such that $|\mathcal{M}| = M \ge P/(R+1)$ with the property that, whenever $x, y \in \mathcal{M}$ with $x \ne y$, one has $||(k!)^k\alpha_{\mathbf{j}}(x-y)|| > P^{1-|\mathbf{j}|}$ for some $\mathbf{j}$ with $2 \le |\mathbf{j}| \le k$. Now recall the numbers $\gamma_{\mathbf{j}}(m)$ defined by (5.2). We may apply the relation (5.6) to deduce that, whenever $x, y \in \mathcal{M}$ with $x \ne y$, there exists $\mathbf{h}$ with $1 \le |\mathbf{h}| \le k-1$ such that $||\gamma_{\mathbf{h}}(x) - \gamma_{\mathbf{h}}(y)|| \gg P^{-|\mathbf{h}|}$. Therefore, by repeating the argument leading to (5.5) in the proof of Theorem 5.1, we may conclude that

$$A^{2s} \ll |f(\boldsymbol{\alpha})|^{2s} \ll M^{-1}(\log P)^{2s}P^{2sd+\Delta},$$

and thus

$$A^{2s} \ll (q_0^{-1} + P^{-1})(\log P)^{2s}QA^{2s} \ll q_0^{-1}(\log P)^{2s}QA^{2s} + A^{2s}P^{-\varepsilon},$$

whence $q_0 \ll Q(\log P)^{2s}$, as required. □

Theorem 5.2 gives us all the information we need to handle the minor arcs for the problem of obtaining an asymptotic formula for $N_{s,k,d}(P)$, since the system (1.2) contains only equations of degree k. In order to obtain the asymptotic formula for $J_{s,k,d}(P)$, however, we need information about rational approximations to the the $\alpha_{\mathbf{j}}$ with $|\mathbf{j}| = 1$ when $|f(\boldsymbol{\alpha})|$ is large, which is not provided by Theorem 5.2. In order to obtain such information, we establish a "final coefficient lemma" analogous to that of Baker [**2**], Lemma 4.6, and this requires us to input some major arc information. We define

$$S(q, \mathbf{a}) = \sum_{\mathbf{x} \in [1,q]^d} e\left(q^{-1} \sum_{1 \le |\mathbf{i}| \le k} a_{\mathbf{i}}\mathbf{x}^{\mathbf{i}}\right), \qquad v(\boldsymbol{\beta}) = \int_{[0,P]^d} e\left(\sum_{1 \le |\mathbf{i}| \le k} \beta_{\mathbf{i}}\boldsymbol{\gamma}^{\mathbf{i}}\right) d\boldsymbol{\gamma},$$

and

$$V(\boldsymbol{\alpha}; q, \mathbf{a}) = q^{-d}S(q, \mathbf{a})v(\boldsymbol{\alpha} - \mathbf{a}/q).$$

The following simple lemma suffices for our purposes.

Lemma 5.3. *One has*

$$f(\boldsymbol{\alpha}) - V(\boldsymbol{\alpha}; q, \mathbf{a}) \ll P^{d-1}\left(q + \sum_{1\le|\mathbf{i}|\le k} |q\alpha_{\mathbf{i}} - a_{\mathbf{i}}|P^{|\mathbf{i}|}\right).$$

Proof. We may clearly suppose that $q \le P$, since otherwise the result is trivial. For each $\mathbf{i}$ with $1 \le |\mathbf{i}| \le k$, we write $\beta_{\mathbf{i}} = \alpha_{\mathbf{i}} - a_{\mathbf{i}}/q$. Sorting into arithmetic progressions modulo q, we obtain

$$f(\boldsymbol{\alpha}) = \sum_{\mathbf{r}\in[1,q]^d} e\left(q^{-1}\sum_{1\le|\mathbf{i}|\le k} a_{\mathbf{i}}\mathbf{r}^{\mathbf{i}}\right)\sum_{\mathbf{j}} e\left(\sum_{1\le|\mathbf{i}|\le k} \beta_{\mathbf{i}}(q\mathbf{j}+\mathbf{r})^{\mathbf{i}}\right),$$

where the second summation is over all $\mathbf{j}$ satisfying $0 \le j_l \le (P-r_l)/q$ for $1 \le l \le d$. By making the change of variables $\boldsymbol{\gamma} = q\mathbf{z} + \mathbf{r}$, we find that

$$v(\boldsymbol{\beta}) = q^d \int_{\mathcal{A}_{\mathbf{r}}} e\left(\sum_{1\le|\mathbf{i}|\le k} \beta_{\mathbf{i}}(q\mathbf{z}+\mathbf{r})^{\mathbf{i}}\right) d\mathbf{z},$$

where

$$\mathcal{A}_{\mathbf{r}} = \{\mathbf{z} \in \mathbb{R}^d : -r_l/q \le z_l \le (P-r_l)/q\}.$$

It follows that

$$f(\boldsymbol{\alpha}) - V(\boldsymbol{\alpha}; q, \mathbf{a}) = \sum_{\mathbf{r}\in[1,q]^d} e\left(q^{-1}\sum_{1\le|\mathbf{i}|\le k} a_{\mathbf{i}}\mathbf{r}^{\mathbf{i}}\right)\left[\sum_{\mathbf{j}}\int_{\mathcal{U}_{\mathbf{j}}} H(\mathbf{z};\mathbf{j},\mathbf{r})\,d\mathbf{z} + O\left((P/q)^{d-1}\right)\right],$$

where

$$H(\mathbf{z};\mathbf{j},\mathbf{r}) = e\left(\sum_{1\le|\mathbf{i}|\le k} \beta_{\mathbf{i}}(q\mathbf{j}+\mathbf{r})^{\mathbf{i}}\right) - e\left(\sum_{1\le|\mathbf{i}|\le k} \beta_{\mathbf{i}}(q\mathbf{z}+\mathbf{r})^{\mathbf{i}}\right)$$

and

$$\mathcal{U}_{\mathbf{j}} = [j_1, j_1+1] \times \cdots \times [j_d, j_d+1].$$

An application of the mean value theorem with respect to the variable z_1 shows that, for $\mathbf{z} \in \mathcal{U}_{\mathbf{j}}$, one has

$$H(\mathbf{z};\mathbf{j},\mathbf{r}) \ll \sum_{1\le|\mathbf{i}|\le k} |\beta_{\mathbf{i}}| q i_1 (qz_1+r_1)^{i_1-1}\cdots(qz_d+r_d)^{i_d} \ll q\sum_{1\le|\mathbf{i}|\le k} |\beta_{\mathbf{i}}|P^{|\mathbf{i}|-1},$$

and the theorem now follows by making trivial estimates. □

We note that a van der Corput analysis along the lines of Baker [**2**], Lemma 4.4, may be applied to give a better error term for small values of q, provided that $|\beta_{\mathbf{i}}| \le (2rkq)^{-1}P^{1-|\mathbf{i}|}$ for each $\mathbf{i}$. Such improvements do not strengthen our final conclusions, however, and we actually find Lemma 5.3 to be more convenient for our purposes.

Before stating our final coefficient lemma, we mention two important estimates. First of all, by Lemma II.2 of [**1**], one has

$$v(\boldsymbol{\beta}) \ll P^d\left(1 + \sum_{1\le|\mathbf{i}|\le k} |\beta_{\mathbf{i}}|P^{|\mathbf{i}|}\right)^{-1/k}. \tag{5.9}$$

Secondly, it follows from Lemma II.8 of [**1**] that, whenever $(q, \mathbf{a}) = 1$, one has

$$S(q, \mathbf{a}) \ll q^{d-1/k+\varepsilon} \tag{5.10}$$

for every $\varepsilon > 0$. These bounds will be used frequently throughout the remainder of our analysis. We are now ready to state the final coefficient lemma.

LEMMA 5.4. *Suppose that $|f(\boldsymbol{\alpha})| \geq A \geq P^{d-\sigma+\varepsilon}$ for some $\varepsilon > 0$, where $\sigma^{-1} > d+1$. Further, write $X = P^{1-(d+1)\sigma}$ and $Y = (P^d A^{-1})^{k+\varepsilon}$, and suppose that there exist integers $v_{\mathbf{j}}$ and w with*

$$1 \leq w \ll X \quad \textit{and} \quad |w\alpha_{\mathbf{j}} - v_{\mathbf{j}}| \ll XP^{-|\mathbf{j}|} \qquad (2 \leq |\mathbf{j}| \leq k).$$

Then there exist integers $a_{\mathbf{j}}$ and q, with $(q, \mathbf{a}) = 1$, such that

$$1 \leq q \ll Y \quad \textit{and} \quad |q\alpha_{\mathbf{j}} - a_{\mathbf{j}}| \ll YP^{-|\mathbf{j}|} \qquad (1 \leq |\mathbf{j}| \leq k).$$

Proof. By Dirichlet's Theorem on simultaneous approximation, we can find an integer t with $1 \leq t \leq P^{\sigma d}$ and integers $a_{\mathbf{j}}$ such that $|tw\alpha_{\mathbf{j}} - a_{\mathbf{j}}| \leq P^{-\sigma}$ for each $\mathbf{j}$ with $|\mathbf{j}| = 1$. Now put $q = tw$, and write $a_{\mathbf{j}} = tv_{\mathbf{j}}$ when $2 \leq |\mathbf{j}| \leq k$. Then we have $q \ll P^{1-\sigma}$ and

$$|q\alpha_{\mathbf{j}} - a_{\mathbf{j}}| = t|w\alpha_{\mathbf{j}} - v_{\mathbf{j}}| \ll P^{1-\sigma-|\mathbf{j}|}.$$

Furthermore, we may divide out common factors to ensure that $(q, \mathbf{a}) = 1$ while preserving the latter two inequalities. Thus by Lemma 5.3 we have

$$P^{d-\sigma+\varepsilon} \leq A \leq |f(\boldsymbol{\alpha})| = |V(\boldsymbol{\alpha}; q, \mathbf{a})| + O(P^{d-\sigma}),$$

and it follows that $A \ll |V(\boldsymbol{\alpha}; q, a)|$. Thus by (5.9) and (5.10), we have

$$A \ll q^{\tau} P^d \biggl(q + \sum_{1 \leq |\mathbf{j}| \leq k} |q\alpha_{\mathbf{j}} - a_{\mathbf{j}}| P^{|\mathbf{j}|} \biggr)^{-1/k},$$

where $\tau = \varepsilon/(2k^2) < 1/(2k)$. It follows that

$$q + \sum_{1 \leq |\mathbf{j}| \leq k} |q\alpha_{\mathbf{j}} - a_{\mathbf{j}}| P^{|\mathbf{j}|} \ll (q^{\tau} P^d A^{-1})^k \ll q^{1/2} (P^d A^{-1})^k.$$

In particular, this shows that $q \ll (P^d A^{-1})^{2k}$ and hence that

$$q + \sum_{1 \leq |\mathbf{j}| \leq k} |q\alpha_{\mathbf{j}} - a_{\mathbf{j}}| P^{|\mathbf{j}|} \ll (P^d A^{-1})^{k+\varepsilon},$$

as required. □

We are now in a position to state the main theorem of this section.

THEOREM 5.5. *Suppose that $|f(\boldsymbol{\alpha})| \geq A \geq P^{d-\sigma+\varepsilon}$ for some $\varepsilon > 0$, where*

$$\sigma^{-1} \geq 2rk(\log k + \tfrac{4}{3} \log r + 2 \log \log k + 3d + 6),$$

and write $Y = (P^d A^{-1})^{k+\varepsilon}$. Then there are integers $a_{\mathbf{j}}$ and q, with $(q, \mathbf{a}) = 1$, satisfying

$$1 \leq q \ll Y \quad \textit{and} \quad |q\alpha_{\mathbf{j}} - a_{\mathbf{j}}| \ll YP^{-|\mathbf{j}|} \qquad (1 \leq |\mathbf{j}| \leq k).$$

Proof. First of all, by applying Theorem 1.1 with

$$s = \lceil rk(\tfrac{1}{2}\log k + \tfrac{2}{3}\log r + \log\log k + 2)\rceil,$$

we find that $\Delta = (\log k)^{-1}$ is an admissible exponent for (s, k, d). Moreover, for fixed s and d, the admissible exponent given by Theorem 1.1 is an increasing function of k, so it follows that Δ is also admissible for $(s, k-1, d)$. A simple calculation reveals that

$$\sigma(4s + d + 1) < 1 - 2\Delta$$

whenever k is sufficiently large, and thus

$$(P^d A^{-1})^{4s} \ll P^{4s\sigma - 2\varepsilon} \ll XP^{-2\Delta - 2\varepsilon},$$

where $X = P^{1-(d+1)\sigma}$. Then by Theorem 5.2, we find that there is an integer q_0 satisfying

$$1 \le q_0 \ll P^{\Delta}(P^d A^{-1})^{2s}(\log P)^{2s} \ll X^{1/2}$$

and

$$||q_0\alpha_{\mathbf{j}}|| \le q_0||q_{\mathbf{j}}\alpha_{\mathbf{j}}|| \ll P^{2\Delta}(P^d A^{-1})^{4s}(\log P)^{2s}P^{-|\mathbf{j}|+\varepsilon} \ll XP^{-|\mathbf{j}|}$$

for all $\mathbf{j}$ with $2 \le |\mathbf{j}| \le k$. Here the integers $q_{\mathbf{j}}$ are as in the statement of Theorem 5.2. We may therefore apply Lemma 5.4 to complete the proof. □

Theorem 1.2 now follows as an easy corollary. Theorem 5.5 is slightly more informative, however, particularly in the situation where $|f(\boldsymbol{\alpha})| \gg P^d$, which arises in the current method for studying diophantine inequalities. In such applications, the fact that q is bounded by a constant is critical.

6. *The asymptotic formulas*

In this section we prove Theorem 1.3 by applying the Hardy-Littlewood method. Essentially the same argument may be applied to deduce Theorem 1.4, and we provide only a sketch of the latter proof.

First recall that

$$J_{s,k,d}(P) = \int_{\mathbb{T}^r} |f(\boldsymbol{\alpha})|^{2s} d\boldsymbol{\alpha}.$$

We let

$$\mathfrak{M}(q, \mathbf{a}) = \{\boldsymbol{\alpha} \in \mathbb{T}^r : |q\alpha_{\mathbf{i}} - a_{\mathbf{i}}| \le P^{1/2-|\mathbf{i}|},\ 1 \le |\mathbf{i}| \le k\} \tag{6.1}$$

and define the set of major arcs $\mathfrak{M}$ to be the union of all $\mathfrak{M}(q, \mathbf{a})$ with $0 \le a_{\mathbf{i}} \le q \le P^{1/2}$ and $(q, \mathbf{a}) = 1$. Further, let $\mathfrak{m} = \mathbb{T}^r \setminus \mathfrak{M}$ denote the minor arcs.

Theorem 6.1. *Let s_1 be as in* (1.13)*, and suppose that $s \ge s_1$. Then there exists $\delta = \delta(k, d) > 0$ such that*

$$\int_{\mathfrak{m}} |f(\boldsymbol{\alpha})|^{2s} d\boldsymbol{\alpha} \ll P^{2sd-K-\delta}.$$

Proof. Write $s = t + u$, where t and u are parameters at our disposal. We have

$$\int_{\mathfrak{m}} |f(\boldsymbol{\alpha})|^{2s} d\boldsymbol{\alpha} \le \sup_{\boldsymbol{\alpha}\in\mathfrak{m}} |f(\boldsymbol{\alpha})|^{2t} \int_{\mathbb{T}^r} |f(\boldsymbol{\alpha})|^{2u} d\boldsymbol{\alpha}.$$

By applying Theorem 1.1, we find that $\Delta_u = (\log k)^{-1}$ is an admissible exponent when

$$u = \lceil rk(\tfrac{2}{3}\log r + \tfrac{1}{2}\log k + \log\log k + 2)\rceil. \tag{6.2}$$

Now suppose that $|f(\boldsymbol{\alpha})| \geq P^{d-\sigma+\varepsilon}$, where $\sigma^{-1} = \frac{8}{3}rk(d+1)\log k \geq \frac{8}{3}rk\log rk$. Then we have $k\sigma \leq 1/2$, so Theorem 1.2 implies that $\boldsymbol{\alpha} \in \mathfrak{M}$. Thus we have

$$\sup_{\boldsymbol{\alpha}\in\mathfrak{m}} |f(\boldsymbol{\alpha})| \ll P^{d-\sigma+\varepsilon}$$

for every $\varepsilon > 0$. It follows that

$$\int_{\mathfrak{m}} |f(\boldsymbol{\alpha})|^{2s} d\boldsymbol{\alpha} \ll P^{2sd-K-2t\sigma+\Delta_u+\varepsilon},$$

and by taking $t > \frac{4}{3}rk(d+1)$ we find that $2t\sigma > \Delta_u$. The proof is now completed by choosing ε sufficiently small in terms of k and d. □

We now write $V(\boldsymbol{\alpha}) = V(\boldsymbol{\alpha}; q, \mathbf{a})$ when $\boldsymbol{\alpha} \in \mathfrak{M}(q,\mathbf{a}) \subseteq \mathfrak{M}$ and define $V(\boldsymbol{\alpha}) = 0$ otherwise. It follows immediately from (6.1) and Lemma 5.3 that

$$f(\boldsymbol{\alpha}) - V(\boldsymbol{\alpha}) \ll P^{d-1/2} \tag{6.3}$$

whenever $\boldsymbol{\alpha} \in \mathfrak{M}$. Moreover, one has

$$\mathrm{meas}(\mathfrak{M}) \ll P^{(r+1)/2-K},$$

and thus

$$\int_{\mathfrak{M}} (|f(\boldsymbol{\alpha})|^{2s} - |V(\boldsymbol{\alpha})|^{2s})\, d\boldsymbol{\alpha} \ll P^{2d-1/2} \int_{\mathfrak{M}} |V(\boldsymbol{\alpha})|^{2s-2} d\boldsymbol{\alpha} + P^{2sd-K-\nu}, \tag{6.4}$$

where $\nu = s - \frac{1}{2}(r+1)$. We are now in a position to handle the major arcs.

THEOREM 6.2. *Whenever $s > \frac{1}{2}k(r+1)+1$, one has*

$$\int_{\mathfrak{M}} |f(\boldsymbol{\alpha})|^{2s} d\boldsymbol{\alpha} = \mathcal{J}\mathfrak{S}P^{2sd-K} + O(P^{2sd-K-\delta}),$$

for some $\delta = \delta(k,d) > 0$, where

$$\mathcal{J} = \int_{\mathbb{R}^r}\int_{[0,1]^{2sd}} e\left(\sum_{1\leq|\mathbf{i}|\leq k} \beta_{\mathbf{i}}(\boldsymbol{\gamma}_1^{\mathbf{i}} + \cdots + \boldsymbol{\gamma}_s^{\mathbf{i}} - \boldsymbol{\gamma}_{s+1}^{\mathbf{i}} - \cdots - \boldsymbol{\gamma}_{2s}^{\mathbf{i}})\right) d\boldsymbol{\gamma}\, d\boldsymbol{\beta}$$

and

$$\mathfrak{S} = \sum_{q=1}^{\infty} \sum_{\substack{\mathbf{a}\in[1,q]^r\\(q,\mathbf{a})=1}} |q^{-d}S(q,\mathbf{a})|^{2s}.$$

Proof. We have

$$\int_{\mathfrak{M}} |V(\boldsymbol{\alpha})|^{2s} d\boldsymbol{\alpha} = \sum_{q\leq P^{1/2}} \sum_{\substack{\mathbf{a}\in[1,q]^r\\(q,\mathbf{a})=1}} |q^{-d}S(q,\mathbf{a})|^{2s} \int_{\mathcal{B}(q)} |v(\boldsymbol{\beta})|^{2s} d\boldsymbol{\beta},$$

where

$$\mathcal{B}(q) = \prod_{1\leq|\mathbf{i}|\leq k} [-q^{-1}P^{1/2-|\mathbf{i}|}, q^{-1}P^{1/2-|\mathbf{i}|}].$$

After two changes of variable, we find that

$$\int_{\mathcal{B}(q)} |v(\boldsymbol{\beta})|^{2s} d\boldsymbol{\beta} = P^{2sd-K}\mathcal{J}(q,P),$$

where

$$\mathcal{J}(q,P) = \int_{\mathcal{B}'(q)} \int_{[0,1]^{2sd}} e\left(\sum_{1\le|\mathbf{i}|\le k} \beta_{\mathbf{i}}(\boldsymbol{\gamma}_1^{\mathbf{i}} + \cdots + \boldsymbol{\gamma}_s^{\mathbf{i}} - \boldsymbol{\gamma}_{s+1}^{\mathbf{i}} - \cdots - \boldsymbol{\gamma}_{2s}^{\mathbf{i}}) \right) d\boldsymbol{\gamma}\, d\boldsymbol{\beta}$$

and $\mathcal{B}'(q) = [-q^{-1}P^{1/2}, q^{-1}P^{1/2}]^r$. Applying (5.9) with $P = 1$ and using the inequality

$$(1 + |\beta_1| + \cdots + |\beta_r|)^r \ge (1 + |\beta_1|)\cdots(1 + |\beta_r|)$$

gives

$$\mathcal{J} - \mathcal{J}(q,P) \ll \int_{q^{-1}P^{1/2}}^{\infty} (1+\beta)^{-2s/rk} d\beta \ll (q^{-1}P^{1/2})^{1-2s/rk}.$$

Combining this with (5.10), we obtain

$$\begin{aligned}\int_{\mathfrak{M}} |V(\boldsymbol{\alpha})|^{2s} d\boldsymbol{\alpha} &= P^{2sd-K} \sum_{q\le P^{1/2}} \sum_{\substack{\mathbf{a}\in[1,q]^r \\ (q,\mathbf{a})=1}} |q^{-d}S(q,\mathbf{a})|^{2s}\mathcal{J}(q,P) \\ &= P^{2sd-K}\left(\mathcal{J} \sum_{q\le P^{1/2}} \sum_{\substack{\mathbf{a}\in[1,q]^r \\ (q,\mathbf{a})=1}} |q^{-d}S(q,\mathbf{a})|^{2s} + E(P) \right),\end{aligned}$$

where

$$E(P) \ll P^{1/2-s/rk} \sum_{q\le P^{1/2}} q^{r-2s/k-1+2s/rk+\varepsilon} \ll P^{-\sigma}$$

for some $\sigma > 0$, since $s > \frac{1}{2}k(r+1)$. In view of (5.10), this lower bound for s also ensures that

$$\sum_{q\le P^{1/2}} \sum_{\substack{\mathbf{a}\in[1,q]^r \\ (q,\mathbf{a})=1}} |q^{-d}S(q,\mathbf{a})|^{2s} = \mathfrak{S} + O(P^{-\tau})$$

for some $\tau > 0$, and we therefore have

$$\int_{\mathfrak{M}} |V(\boldsymbol{\alpha})|^{2s} d\boldsymbol{\alpha} = \mathcal{J}(\mathfrak{S} + O(P^{-\delta}))P^{2sd-K}, \tag{6.5}$$

where we have set $\delta = \min(\sigma,\tau)$. Moreover, since $s - 1 > \frac{1}{2}k(r+1)$, we see that (6.5) also holds with s replaced by $s-1$. The theorem now follows on recalling (6.4), since (5.9) implies that $\mathcal{J} \ll 1$. $\square$

The proof of Theorem 1.3 is now completed by combining Theorems 6.1 and 6.2 and noting that, in view of (1.9), one has $\mathcal{J}\mathfrak{S} > 0$.

If one is prepared to suppose the existence of non-singular real and p-adic solutions to the system (1.2), then the methods illustrated above can be used to establish an asymptotic formula for $N_{s,k,d}(P)$ whenever $s \ge 2s_1$, as claimed in the statement of Theorem 1.4. We provide only a brief sketch of the argument. First of all, one has

$$N_{s,k,d}(P) = \int_{\mathbb{T}^\ell} \left(\prod_{j=1}^{s} f_j(\boldsymbol{\alpha}) \right) d\boldsymbol{\alpha}, \tag{6.6}$$

where we have written $\ell = \binom{k+d-1}{k}$ for the number of equations in (1.2), and where

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

We define $\mathfrak{M}(q, \mathbf{a})$ as in (6.1), except that the condition $1 \le |\mathbf{i}| \le k$ is replaced by $|\mathbf{i}| = k$, and we again take $\mathfrak{M}$ to be the union of the $\mathfrak{M}(q, \mathbf{a})$ with $0 \le a_{\mathbf{i}} \le q \le P^{1/2}$ and $(q, \mathbf{a}) = 1$. Next we write $s = t + 2u$, where $t > \frac{8}{3} rk(d+1)$ and u is as in (6.2). After applying Hölder's inequality and making a change of variable, we may apply (1.15) to conclude as in the proof of Theorem 6.1 that the minor arc contribution to the integral (6.6) is of order at most $P^{sd-L-\nu}$ for some $\nu > 0$.

Furthermore, by repeating the argument of the proof of Theorem 6.2, one finds that

$$\int_{\mathfrak{M}} \left(\prod_{j=1}^{s} f_j(\boldsymbol{\alpha}) \right) d\boldsymbol{\alpha} = \mathcal{J}_1 \mathfrak{S}_1 P^{sd-L} + O(P^{sd-L-\nu}),$$

for some $\nu > 0$, where

$$\mathcal{J}_1 = \int_{\mathbb{R}^\ell} \int_{[-1,1]^{sd}} e\left(\sum_{|\mathbf{i}|=k} \beta_{\mathbf{i}} (c_1 \boldsymbol{\gamma}_1^{\mathbf{i}} + \cdots + c_s \boldsymbol{\gamma}_s^{\mathbf{i}}) \right) d\boldsymbol{\gamma}\, d\boldsymbol{\beta}$$

and

$$\mathfrak{S}_1 = \sum_{q=1}^{\infty} \sum_{\substack{\mathbf{a} \in [1,q]^\ell \\ (q,\mathbf{a})=1}} \prod_{j=1}^{s} \left(q^{-d} S(q, c_j \mathbf{a}) \right).$$

To show that $\mathcal{J}_1 > 0$, one does some analysis in a neighborhood of the non-singular real solution as in the argument of [**8**], Lemma 7.4, for example. To show that $\mathfrak{S}_1 > 0$, one makes use of the non-singular p-adic solutions within a Hensel's lemma-type argument (see for instance [**7**], Lemmas 9.6–9.9).

Acknowledgements. The author expresses his thanks to Professors Vaughan and Wooley for years of encouragement and helpful discussions. The author is also grateful to the referees of this paper for many insightful comments.

References

1. G. I. Arkhipov, A. A. Karatsuba and V. N. Chubarikov, 'Multiple trigonometric sums', *Trudy Mat. Inst. Steklov* 151 (1980) 1–126.

2. R. C. Baker, *Diophantine inequalities* (Clarendon Press, Oxford, 1986).

3. B. J. Birch, 'Homogeneous forms of odd degree in a large number of variables', *Mathematika* 4 (1957) 102–105.

4. R. Brauer, 'A note on systems of homogeneous algebraic equations', *Bull. Amer. Math. Soc.* 51 (1945) 749–755.

5. K. B. Ford, 'New estimates for mean values of Weyl sums', *Internat. Math. Res. Notices* (1995) 155–171.

6. S. T. Parsell, 'The density of rational lines on cubic hypersurfaces', *Trans. Amer. Math. Soc.* 352 (2000) 5045–5062.

7. S. T. Parsell, 'Multiple exponential sums over smooth numbers', *J. Reine Angew. Math.* 532 (2001) 47–104.

8. S. T. Parsell, 'Pairs of additive equations of small degree', *Acta Arith.* 104 (2002) 345–402.

9. R. C. Vaughan, *The Hardy-Littlewood method*, 2nd edn (Cambridge University Press, Cambridge, 1997).

10. I. M. Vinogradov, 'The method of trigonometrical sums in the theory of numbers', *Trav. Inst. Skeklov.* 23 (1947).

11. T. D. Wooley, 'Large improvements in Waring's problem', *Ann. of Math.* (2) 135 (1992) 131–164.

12. T. D. Wooley, 'On Vinogradov's mean value theorem', *Mathematika* 39 (1992) 379–399.

13. T. D. Wooley, 'A note on symmetric diagonal equations', *Number Theory with an emphasis on the Markoff spectrum* (ed. A. D. Pollington and W. Moran, Marcel Dekker, 1993), pp. 317–321.

14. T. D. Wooley, 'A note on simultaneous congruences', *J. Number Theory* 58 (1996) 288–297.

15. T. D. Wooley, 'Some remarks on Vinogradov's mean value theorem and Tarry's problem', *Monatsh. Math.* 122 (1996) 265–273.

16. T. D. Wooley, 'On exponential sums over smooth numbers', *J. Reine Angew. Math.* 488 (1997) 79–140.

Scott T. Parsell
Department of Mathematics and Actuarial Science
Butler University
4600 Sunset Avenue, JH 270
Indianapolis, IN 46208
U.S.A.

sparsell@butler.edu