MULTIPLE EXPONENTIAL SUMS OVER SMOOTH NUMBERS

SCOTT T. PARSELL

ABSTRACT. We obtain estimates for mean values of double exponential sums over smooth numbers by developing a suitable version of the Vaughan-Wooley iterative method. These estimates are then used within the fabric of the Hardy-Littlewood method to provide a lower bound for the density of rational lines on the hypersurface defined by an additive equation when the dimension is sufficiently large in terms of the degree. We also consider applications to a two-dimensional generalization of Waring's problem.

1. INTRODUCTION

Let $F(\mathbf{x})$ be a form of degree k in s variables, with integer coefficients. In 1945, Brauer [4] demonstrated the existence of an m-dimensional linear space on the hypersurface $F(\mathbf{x}) = 0$ over some solvable extension of $\mathbb{Q}$, provided that s is sufficiently large in terms of k and m, and in 1957 Birch [3] obtained the same result over $\mathbb{Q}$ for odd k. Unfortunately, the elementary methods of Brauer and Birch do not yield any reasonable bounds on the number of variables required, although explicit calculations have been done more recently for small values of k by Lewis and Schulze-Pillot [8] and Wooley [15], [16]. Moreover, up to this point no estimates have been provided for the density of rational lines on a given hypersurface.

In this paper, we obtain an explicit upper bound for the number of variables required to guarantee the expected density of lines on the hypersurface $F(\mathbf{x}) = 0$ in the case when F is an additive form of degree k. Our approach is via the Hardy-Littlewood method, and we will be required to develop considerable analytic machinery in order to get started. The method depends fundamentally on sharp estimates for certain multiple exponential sums over smooth numbers, which we obtain by extending the ideas of Vaughan [11] and Wooley [13], [17]. Such estimates are of interest in their own right and may also be applied, for example, to the two-dimensional generalization of Waring's problem proposed by Arkhipov and Karatsuba [1], which we consider in Section 9.

When P and R are positive integers, write

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

for the set of R-smooth numbers up to P, and define the exponential sum

$$f(\boldsymbol{\alpha}; P, R) = \sum_{x,y \in \mathcal{A}(P,R)} e(\alpha_0 x^k + \alpha_1 x^{k-1} y + \cdots + \alpha_k y^k). \tag{1.1}$$

Further, define the mean value

$$S_s(P,R) = \int_{\mathbb{T}^{k+1}} |f(\boldsymbol{\alpha}; P, R)|^{2s} d\boldsymbol{\alpha},$$

1991 *Mathematics Subject Classification.* 11L03, 11L07, 11L15, 11P05, 11P55, 11D85.

Research supported in part by NSF grant DMS-9622773 and by a fellowship from the David and Lucile Packard Foundation.

and observe that $S_s(P,R)$ is the number of solutions of the system of equations

$$\sum_{m=1}^{s}(x_m^{k-i}y_m^i - \tilde{x}_m^{k-i}\tilde{y}_m^i) = 0 \quad (0 \le i \le k) \tag{1.2}$$

with

$$x_m, y_m, \tilde{x}_m, \tilde{y}_m \in \mathcal{A}(P,R) \quad (1 \le m \le s).$$

The following theorem provides a simple upper bound for $S_s(P,R)$.

Theorem 1. *Let $k \ge 2$ be a positive integer, and put $r = \left[\frac{k+1}{2}\right]$. Further, write*

$$s_1 = k^2\left(1 - \frac{1}{2k}\right)^{-1} + r,$$

and let s be a positive integer with $s \ge s_1$. Then for any $\varepsilon > 0$ there exists $\eta = \eta(s,k,\varepsilon)$ such that

$$S_s(P,R) \ll P^{4s-k(k+1)+\Delta_s+\varepsilon}, \tag{1.3}$$

where $R \le P^\eta$ and

$$\Delta_s = k(k+1)\left(1 - \frac{1}{2k}\right)^{(s-s_1)/r}.$$

For example, if $s \sim 2k^2(\log k + \log\log k)$, then we have $\Delta_s \ll k^2 e^{-s/k^2} \ll (\log k)^{-2}$. Whenever Δ_s has the property that, for every $\varepsilon > 0$, there exists $\eta = \eta(\varepsilon)$ such that (1.3) holds whenever $R \le P^\eta$, we say that Δ_s is an admissible exponent.

We note for comparison that Arkhipov, Karatsuba, and Chubarikov [2] have obtained estimates for the number of solutions of the "complete" system

$$\sum_{m=1}^{s}(x_m^i y_m^j - \tilde{x}_m^i\tilde{y}_m^j) = 0 \quad (0 \le i,j \le k)$$

with

$$1 \le x_m, y_m, \tilde{x}_m, \tilde{y}_m \le P \quad (1 \le m \le s)$$

which lead, via a standard argument, to admissible exponents for (1.2) behaving roughly like $k^3 e^{-s/2k^3}$, so that one must take $s \ge 6k^3\log k$ in most applications.

We also remark that, when $R = P^\eta$, an elementary argument yields the lower bound

$$S_s(P,R) \gg P^{2s} + P^{4s-k(k+1)} \tag{1.4}$$

and that a weak upper bound of the form

$$S_s(P,R) \ll P^{4s-\frac{1}{2}k(k+1)+\Delta_s'+\varepsilon},$$

follows on fixing $\mathbf{y}, \tilde{\mathbf{y}}$ and applying the results of [17] to the equations in $\mathbf{x}, \tilde{\mathbf{x}}$.

In Section 6, we obtain the following sharper result as a consequence of repeated efficient differencing.

Theorem 2. *Write* $r = \left[\frac{k+1}{2}\right]$, *and put*

$$s_0 = k(k+1) \qquad \text{and} \qquad s_1 = \tfrac{4}{3} rk(\log(4rk) - 2\log\log k).$$

Further, define

$$\Delta_s = \begin{cases} 4rke^{2-3(s-s_0)/4rk}, & \text{when } 1 \le s \le s_1, \\ e^4(\log k)^2\left(1-\frac{1}{2k}\right)^{(s-s_1)/r}, & \text{when } s > s_1. \end{cases}$$

Then there exists a constant K such that the exponent Δ_s is admissible whenever $k \ge K$.

Notice that the admissible exponents one obtains from Theorem 2 decay in most cases roughly like $k^2 e^{-3s/2k^2}$, whereas those obtained from Theorem 1 decay like $k^2 e^{-s/k^2}$.

The mean value estimates of Theorems 1 and 2 may be transformed into Weyl estimates by using the large sieve inequality in a standard way. Thus in Section 7 we will prove the following result.

Theorem 3. *For $\mu > 0$, define $\mathfrak{m}_\mu$ to be the set of $\boldsymbol{\alpha} \in \mathbb{R}^{k+1}$ such that whenever $a_i \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $(a_0, \dots, a_k, q) = 1$ and $|q\alpha_i - a_i| \le P^{\mu-k}R^k$ $(0 \le i \le k)$ one has $q > P^\mu R^{k+1}$. Suppose that $0 < \lambda \le \frac{1}{2}$ and that Δ_s denotes an admissible exponent. Then given $\varepsilon > 0$ there exists $\eta = \eta(\varepsilon, k)$ such that whenever $R \le P^\eta$ one has*

$$\sup_{\boldsymbol{\alpha}\in\mathfrak{m}_{\lambda(k+1)}} |f(\boldsymbol{\alpha}; P, R)| \ll P^{2-\sigma(\lambda)+\varepsilon},$$

where

$$\sigma(\lambda) = \max_{2s \ge k+1} \frac{\lambda - (1-\lambda)\Delta_s}{2s}. \tag{1.5}$$

In our applications involving the circle method, we will find it useful to take $\lambda = \frac{1}{2(k+1)}$. After performing a simple optimization, one obtains the following simplification.

Corollary 3.1. *Given $\varepsilon > 0$, there exists $\eta = \eta(\varepsilon, k)$ such that whenever $R \le P^\eta$ one has*

$$\sup_{\boldsymbol{\alpha}\in\mathfrak{m}_{1/2}} |f(\boldsymbol{\alpha}; P, R)| \ll P^{2-\sigma_1(k)+\varepsilon},$$

where

$$\sigma_1(k)^{-1} \sim \tfrac{28}{3} k^3 \log k$$

as $k \to \infty$.

We now consider the multidimensional analogue of Waring's problem discussed in [1]. Let $W_s(\mathbf{n}, P)$ denote the number of solutions of the system of equations

$$x_1^{k-j}y_1^j + \dots + x_s^{k-j}y_s^j = n_j \quad (0 \le j \le k) \tag{1.6}$$

with $x_i, y_i \in [1, P] \cap \mathbb{Z}$. Obviously, $W_s(\mathbf{n}, P) = 0$ if the relative sizes of the n_j are incompatible, since then the equations (1.6) will be insoluble even over the positive reals. Thus we will need to impose some conditions in order to proceed.

Theorem 4. *Suppose that*

$$s \ge \tfrac{14}{3}k^2\log k + \tfrac{10}{3}k^2\log\log k + O(k^2),$$

and fix real numbers $\mu_0, \dots, \mu_k$ with the property that the system

$$\eta_1^{k-j}\xi_1^j + \cdots + \eta_s^{k-j}\xi_s^j = \mu_j \quad (0 \le j \le k) \tag{1.7}$$

has a non-singular real solution with $0 < \eta_i, \xi_i < 1$. Suppose also that the system (1.6) has a non-singular p-adic solution for all primes p. Then there exist positive numbers $P_0 = P_0(s, k, \boldsymbol{\mu})$ and $\delta = \delta(s, k, \boldsymbol{\mu})$ such that, whenever $P > P_0$ and

$$|n_j - P^k\mu_j| < \delta P^k \quad (0 \le j \le k), \tag{1.8}$$

one has

$$W_s(\mathbf{n}, P) \gg P^{2s-k(k+1)}.$$

We remark that the p-adic solubility condition imposed in the theorem in fact need only be checked for finitely many primes p, as we will see in Section 9 that primes sufficiently large in terms of k may be dealt with unconditionally using exponential sums.

Theorem 4 leads, via the binomial theorem, to the conclusion that suitable polynomials of degree k with integer coefficients may be represented as sums of kth powers of linear polynomials. That is, we seek to write

$$p(t) = (x_1 t + y_1)^k + \cdots + (x_s t + y_s)^k \tag{1.9}$$

with $x_i, y_i \in \mathbb{N}$.

We will say that the polynomial

$$p(t) = \sum_{j=0}^{k} \binom{k}{j} n_j t^j \tag{1.10}$$

is locally representable if

(1) there exist real numbers P, δ, and $\mu_0, \dots, \mu_k$ such that (1.8) holds and such that the system (1.7) has a non-singular real solution with $0 < \eta_i, \xi_i < 1$, and
(2) the system (1.6) has a non-singular p-adic solution for all primes p.

Now let $G_1^*(k)$ denote the least integer s such that, whenever the polynomial $p(t)$ given by (1.10) is locally representable and $n_0, \dots, n_k$ are sufficiently large, one has the global representation (1.9) for some natural numbers $x_1, \dots, x_s$ and $y_1, \dots, y_s$.

From Theorem 4 we immediately obtain an upper bound for $G_1^*(k)$.

Corollary 4.1. *One has*

$$G_1^*(k) \le \tfrac{14}{3}k^2 \log k + \tfrac{10}{3}k^2 \log\log k + O(k^2).$$

We note that Arkhipov and Karatsuba [1] have previously outlined a program for obtaining bounds of the form $G_1^*(k) \le Ck^2 \log k$ using the theory of multiple exponential sums over a complete interval developed in [2]. Corollary 4.1 thus gives an explicit asymptotic version of this result, showing that one may take $C \sim 14/3$.

It is worth noting that the analogous problem over the complex numbers has been considered recently by algebraic geometers (see for example [7], [9]). By exploiting a surprising connection with the theory of partial differential operators, one finds that precisely $s = \left\lceil \frac{k+1}{2} \right\rceil$ terms are required to guarantee a representation of the shape (1.9) for arbitrary polynomials of degree k over $\mathbb{C}$. In fact, similar results are known when $p(t)$ is replaced by a form in several variables.

Finally, we return to the problem posed at the beginning of the paper, namely, counting rational lines on the hypersurface defined by an additive equation. Let $c_1, \dots, c_s$ be nonzero integers, and write $N_s(P)$ for the number of solutions of the polynomial equation

$$c_1(x_1t + y_1)^k + \cdots + c_s(x_st + y_s)^k = 0 \tag{1.11}$$

with $x_i, y_i \in [-P, P] \cap \mathbb{Z}$. Equivalently, by the binomial theorem, $N_s(P)$ is the number of solutions of the system of equations

$$c_1x_1^{k-j}y_1^j + \cdots + c_sx_s^{k-j}y_s^j = 0 \quad (0 \le j \le k) \tag{1.12}$$

with $x_i, y_i \in [-P, P] \cap \mathbb{Z}$.

Theorem 5. *Suppose that*

$$s \ge \tfrac{14}{3}k^2 \log k + \tfrac{10}{3}k^2 \log\log k + O(k^2),$$

and that the system of equations (1.12) *has a non-singular real solution and a non-singular p-adic solution for all primes p. Then for P sufficiently large one has*

$$N_s(P) \gg P^{2s-k(k+1)}.$$

As in Theorem 4, the p-adic solubility hypothesis here need only be verified for small primes, as the primes $p > p_0(k)$ are easily dealt with by an analytic argument.

Given a line $\ell : \mathbf{x}t + \mathbf{y}$, we define the height of ℓ by $h(\ell) = \max(|x_i|, |y_i|)$. To obtain the density result mentioned in the opening, we seek a lower bound for the number of lines ℓ on our hypersurface that satisfy $h(\ell) \le P$. Among the solutions counted by $N_s(P)$, we may of course have several that correspond to the same line, so Theorem 5 does not directly yield such a lower bound. In Section 10, however, we will actually derive the estimate of Theorem 5 when the variables are restricted to lie in dyadic-type intervals and then show that in this situation the number of solutions of (1.12) corresponding to any particular line is at most $O(1)$. Thus we will prove the following theorem.

Theorem 6. *Let $L_s(P)$ denote the number of distinct rational lines ℓ lying on the hypersurface*

$$c_1z_1^k + \cdots + c_sz_s^k = 0 \tag{1.13}$$

and satisfying $h(\ell) \le P$. Then, under the hypotheses of Theorem 5, *one has*

$$L_s(P) \gg P^{2s-k(k+1)}.$$

We note that, when s is large in terms of k, the theory of a single additive equation (see for example [12]) shows that the hypersurface defined by (1.13) contains "trivial" lines, corresponding to the case where either $x_i = 0$ or $y_i = 0$ for each i in (1.11). By a trivial estimate, however, the number of such lines is $O(P^s)$. Hence Theorem 6 shows that in this situation most of the points on (1.13) that lie on lines in fact lie on non-trivial lines.

For a hypersurface defined by an additive cubic equation, the author's forthcoming work [10] shows that the estimate of Theorem 6 holds unconditionally whenever $s \ge 57$.

The author wishes to thank Professor Trevor Wooley for suggesting these problems and for providing substantial advice and encouragement during the writing of this paper. The author also acknowledges the important comments of the referee.

2. Preliminary Lemmata

Before embarking on the proofs of our mean value estimates, we need to make some preliminary observations. We start by showing that solutions of (1.2) in which some x_j and y_j or some $\tilde{x}_j$ and $\tilde{y}_j$ have a large common factor can effectively be ignored. When $\gamma > 0$, let $S_s(P, R; \gamma)$ be the number of solutions of (1.2) with $(x_j, y_j) \le P^\gamma$ and $(\tilde{x}_j, \tilde{y}_j) \le P^\gamma$ for all j.

Lemma 2.1. *For every $\gamma > 0$, one has $S_s(P, R) \ll P^{2s+\varepsilon} + S_s(P, R; \gamma)$.*

Proof. Write $S'_s(P, R; \gamma)$ for the number of solutions of (1.2) with $(x_j, y_j) > P^\gamma$ or $(\tilde{x}_j, \tilde{y}_j) > P^\gamma$ for some j, so that $S_s(P, R) = S_s(P, R; \gamma) + S'_s(P, R; \gamma)$. Then we have

$$S'_s(P, R; \gamma) = \sum_{d > P^\gamma} \int_{\mathbb{T}^{k+1}} f(d^k \boldsymbol{\alpha}; P/d, R) f(-\boldsymbol{\alpha}; P, R) |f(\boldsymbol{\alpha}; P, R)|^{2s-2} \, d\boldsymbol{\alpha}. \tag{2.1}$$

Now suppose that $S'_s(P, R; \gamma) \ge S_s(P, R; \gamma)$, so that $S_s(P, R) \le 2S'_s(P, R; \gamma)$, and let

$$\lambda_s = \inf\{\lambda : S_s(P, R) \ll P^\lambda\}.$$

If $\lambda_s = 2s$, then we are done, so in view of (1.4) we may assume that $\lambda_s > 2s$. By applying Hölder's inequality to (2.1), we obtain

$$S_s(P, R) \ll \sum_{d > P^\gamma} \left(\int_{\mathbb{T}^{k+1}} |f(d^k \boldsymbol{\alpha}; P/d, R)|^{2s} \, d\boldsymbol{\alpha} \right)^{1/2s} \left(\int_{\mathbb{T}^{k+1}} |f(\boldsymbol{\alpha}; P, R)|^{2s} \, d\boldsymbol{\alpha} \right)^{1-1/2s},$$

from which we deduce that

$$S_s(P, R) \ll \left(\sum_{d > P^\gamma} S_s(P/d, R)^{1/2s} \right)^{2s} \ll P^{\lambda_s + \gamma(2s - \lambda_s) + \varepsilon},$$

for all $\varepsilon > 0$, since $\lambda_s > 2s$. This provides a contradiction for ε sufficiently small, so in fact we have $S'_s(P, R; \gamma) < S_s(P, R; \gamma)$, and the conclusion of the lemma follows. $\square$

We next record an estimate for the number of solutions of an associated system of congruences. When $f_1, \ldots, f_t$ are polynomials in $\mathbb{Z}[x_1, \ldots, x_t]$, write $\mathcal{B}_t(q, p; \mathbf{u}; \mathbf{f})$ for the set of solutions modulo $q^k p^k$ of the simultaneous congruences

$$f_j(x_1, \ldots, x_t) \equiv u_j \pmod{q^{k-j+1} p^{j-1}} \quad (1 \le j \le t) \tag{2.2}$$

with $(J_t(\mathbf{f}; \mathbf{x}), pq) = 1$, where

$$J_t(\mathbf{f}; \mathbf{x}) = \det \left(\frac{\partial f_i}{\partial x_j}(\mathbf{x}) \right)_{1 \le i, j \le t}. \tag{2.3}$$

Lemma 2.2. *Suppose that $f_1, \ldots, f_{2r} \in \mathbb{Z}[x_1, \ldots, x_{2r}]$ have degrees bounded in terms of k. Then whenever $2r \le k + 1$ we have*

$$\operatorname{card}(\mathcal{B}_{2r}(q, p; \mathbf{u}; \mathbf{f})) \ll_{\varepsilon, k} (pq)^{r(2r-1)+\varepsilon} (q, p)^{2r(2k-2r+1)}.$$

Proof. Write $\tilde{q} = q/(q, p)$ and $\tilde{p} = p/(q, p)$, so that $(\tilde{q}, \tilde{p}) = 1$. Then by considering the jth congruence in (2.2) modulo $\tilde{q}^{k-j+1}$, we obtain from Lemma 2.2 of Wooley [17] that the number of solutions modulo $\tilde{q}^k$ is $O_{\varepsilon,k}(\tilde{q}^{r(2r-1)+\varepsilon})$. Similarly, the number of solutions modulo $\tilde{p}^k$ is $O_{\varepsilon,k}(\tilde{p}^{r(2r-1)+\varepsilon})$. Hence by the Chinese Remainder Theorem the number of solutions modulo $\tilde{q}^k\tilde{p}^k$ is $O_{\varepsilon,k}((\tilde{p}\tilde{q})^{r(2r-1)+\varepsilon})$. Trivially, each of these solutions lifts in at most $(q,p)^{4kr}$ ways to $\mathbb{Z}/(q^kp^k)$, and the lemma follows immediately. □

We now develop some notation for analyzing real singular solutions of systems such as (1.2). Let $\psi_1,\dots,\psi_{2r}$ be nontrivial polynomials in $\mathbb{Z}[x,y]$ of total degree at most k. When $\mathcal{I},\mathcal{J}\subset\{1,2,\dots,2r\}$ with $\mathrm{card}(\mathcal{J})=2\,\mathrm{card}(\mathcal{I})$ and $\mathbf{z},\mathbf{w}\in\mathbb{Z}^{2r}$, define the Jacobian

$$J(\mathcal{I},\mathcal{J};\boldsymbol{\psi})=\det\begin{pmatrix}\frac{\partial\psi_j}{\partial z_i}(z_i,w_i)\\ \frac{\partial\psi_j}{\partial w_i}(z_i,w_i)\end{pmatrix}_{i\in\mathcal{I},j\in\mathcal{J}}.$$

Write $\mathcal{J}_d=\{1,\dots,d\}$, and let $\mathcal{I}_d^*$ denote the set of all subsets of $\mathcal{J}_{2r}$ of size d. We will call the $4r$-tuple of integers $(z_1,w_1,\dots,z_{2r},w_{2r})$ highly singular for $\boldsymbol{\psi}$ if $J(\mathcal{I},\mathcal{J}_{2r};\boldsymbol{\psi})=0$ for each $\mathcal{I}\in\mathcal{I}_r^*$. Also write

$$d_{i,j}(z,w;\boldsymbol{\psi})=\det\begin{pmatrix}\frac{\partial\psi_i}{\partial z}(z,w) & \frac{\partial\psi_j}{\partial z}(z,w)\\ \frac{\partial\psi_i}{\partial w}(z,w) & \frac{\partial\psi_j}{\partial w}(z,w)\end{pmatrix}.$$

Let $\mathcal{S}_r(P;\boldsymbol{\psi})$ denote the set of all integral $4r$-tuples $(z_1,w_1,\dots,z_{2r},w_{2r})$ with $1\le z_i,w_i\le P$ which are highly singular for $\boldsymbol{\psi}$.

Lemma 2.3. *Suppose that $\psi_1,\dots,\psi_{2r}$ satisfy the condition that $d_{1,2}$ is non-trivial and $\deg_w(d_{i,j})<\deg_w(d_{i',j'})$ whenever $i+j<i'+j'$. Then we have*

$$\mathrm{card}(\mathcal{S}_r(P;\boldsymbol{\psi}))\ll_k P^{3r-1}.$$

Proof. Let $\mathcal{T}_0(P;\boldsymbol{\psi})$ denote the set of integral $4r$-tuples $(\mathbf{z},\mathbf{w})$ with $1\le z_i,w_i\le P$ for $i=1,\dots,2r$ and

$$d_{1,2}(z_i,w_i;\boldsymbol{\psi})=0 \tag{2.4}$$

for all i. For a $4r$-tuple counted by $\mathcal{T}_0(P;\boldsymbol{\psi})$ and a given i, there are at most $O(P)$ choices for z_i and w_i satisfying (2.4), since we have assumed that $d_{1,2}$ is non-trivial, and it follows that $\mathrm{card}(\mathcal{T}_0(P;\boldsymbol{\psi}))\ll P^{2r}$.

Now for $1\le d\le r-1$, we say that $(\mathbf{z},\mathbf{w})\in\mathcal{T}_d(P;\boldsymbol{\psi})$ if

$$J(\mathcal{I},\mathcal{J}_{2d};\boldsymbol{\psi})\neq 0 \tag{2.5}$$

for some $\mathcal{I}\in\mathcal{I}_d^*$ but

$$J(\mathcal{I}\cup\{i\},\mathcal{J}_{2d+2};\boldsymbol{\psi})=0 \tag{2.6}$$

for all $i\in\mathcal{J}_{2r}\setminus\mathcal{I}$. Consider a $4r$-tuple counted by $\mathcal{T}_d(P;\boldsymbol{\psi})$ for some $1\le d\le r-1$. There are $O(1)$ choices for $\mathcal{I}$ and $O(P^{2d})$ choices for the z_i and w_i with $i\in\mathcal{I}$. Now we fix $i\in\mathcal{J}_{2r}\setminus\mathcal{I}$ and expand the determinant in (2.6) using 2×2 blocks along the rows containing z_i and w_i. Then on using (2.5) together with our hypothesis on $\boldsymbol{\psi}$, we see that the relation (2.6) is a non-trivial polynomial equation in the variables z_i and w_i and hence has $O(P)$ solutions. Thus we have

$$\mathrm{card}(\mathcal{T}_d(P;\boldsymbol{\psi}))\ll P^{2d+(2r-d)}=P^{2r+d}$$

and hence

$$\mathrm{card}(\mathcal{S}_r(P;\boldsymbol{\psi}))\le\sum_{d=0}^{r-1}\mathrm{card}(\mathcal{T}_d(P;\boldsymbol{\psi}))\ll P^{3r-1},$$

as desired. □

Finally, we recall an estimate of Wooley [14] for the number of integers in an interval with a given square-free kernel. We adopt the notation $s_0(N) = \prod_{p|N} p$.

Lemma 2.4. *Suppose that L is a positive real number and r is a positive integer with $\log r \ll \log L$. Then for each $\varepsilon > 0$, one has*

$$\operatorname{card}\{y \le L : s_0(y) = s_0(r)\} \ll_\varepsilon L^\varepsilon.$$

Proof. This is Lemma 2.1 of Wooley [14]. ☐

3. The Fundamental Lemma

For $0 \le i \le k$, let $\psi_i(z, w; \mathbf{c})$ be polynomials with integer coefficients in the variables $z, w, c_1, \dots, c_u$ and satisfying the conditions of Lemma 2.3. Further, suppose that C_i and C_i' satisfy $1 \le C_i' \le C_i \ll P$, write

$$\tilde{C} = \prod_{i=1}^{u} C_i,$$

and let $D_i(\mathbf{c})$ be polynomials with total degrees bounded in terms of k such that $D_i(\mathbf{c}) \ne 0$ for $C_i' \le c_i \le C_i$. We let ε, η, and γ denote small positive numbers, whose values may change from statement to statement. Generally, η and γ will be chosen sufficiently small in terms of ε, and the implicit constants in our analysis may depend at most on ε, η, γ, s, and k. Since our methods will involve only a finite number of steps, all implicit constants that arise remain under control, and the values assumed by η and γ throughout the arguments remain uniformly bounded away from zero.

When $r \le \left[\frac{k+1}{2}\right]$, let $S_{s,r}(P, Q, R; \boldsymbol{\psi}) = S_{s,r}(P, Q, R; \boldsymbol{\psi}; \mathbf{C}, \mathbf{D}; \gamma)$ be the number of solutions of the system

$$\begin{aligned}\sum_{n=1}^{r} \eta_n(\psi_i(z_n, w_n; \mathbf{c}) - \psi_i(\tilde{z}_n, \tilde{w}_n; \mathbf{c}))& \\ + D_i(\mathbf{c}) \sum_{m=1}^{s} (x_m^{k-i} y_m^i - \tilde{x}_m^{k-i} \tilde{y}_m^i) = 0 \quad (0 \le i \le k)&\end{aligned} \tag{3.1}$$

with

$$x_m, y_m, \tilde{x}_m, \tilde{y}_m \in \mathcal{A}(Q, R) \quad (1 \le m \le s), \tag{3.2}$$

$$(x_m, y_m) \le P^\gamma \quad \text{and} \quad (\tilde{x}_m, \tilde{y}_m) \le P^\gamma \quad (1 \le m \le s), \tag{3.3}$$

$$1 \le z_n, w_n, \tilde{z}_n, \tilde{w}_n \le P \quad \text{and} \quad \eta_n \in \{\pm 1\} \quad (1 \le n \le r), \tag{3.4}$$

and

$$C_j' \le c_j \le C_j \quad (1 \le j \le u). \tag{3.5}$$

Further, write $\tilde{S}_{s,r}(P, Q, R; \boldsymbol{\psi})$ for the number of solutions of (3.1) with (3.2), (3.3), (3.4), (3.5), and

$$J_{2r}(\mathbf{z}, \mathbf{w}; \mathbf{c}) \ne 0 \quad \text{and} \quad J_{2r}(\tilde{\mathbf{z}}, \tilde{\mathbf{w}}; \mathbf{c}) \ne 0, \tag{3.6}$$

where (recalling the notation of the previous section) we have put

$$J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c}) = J(\mathcal{J}_r, \mathcal{J}_{2r}, \boldsymbol{\psi}(\mathbf{z},\mathbf{w};\mathbf{c})).$$

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

$$\begin{aligned}\sum_{n=1}^{r} \eta_n(\psi_i(z_n,w_n;\mathbf{c}) - \psi_i(\tilde{z}_n,\tilde{w}_n;\mathbf{c}))\\ + D_i(\mathbf{c})q^{k-i}p^i \sum_{m=1}^{s} (u_m^{k-i}v_m^i - \tilde{u}_m^{k-i}\tilde{v}_m^i) = 0 \quad (0 \le i \le k)\end{aligned} \tag{3.7}$$

with (3.4), (3.5),

$$P^\theta < p, q \le P^\theta R \quad \text{and} \quad (q,p) \le P^\gamma, \tag{3.8}$$

$$u_m, v_m, \tilde{u}_m, \tilde{v}_m \in \mathcal{A}(QP^{-\theta}, R) \quad (1 \le m \le s), \tag{3.9}$$

$$(u_m, v_m) \le P^\gamma \quad \text{and} \quad (\tilde{u}_m, \tilde{v}_m) \le P^\gamma \quad (1 \le m \le s), \tag{3.10}$$

and

$$(J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c}), pq) = (J_{2r}(\tilde{\mathbf{z}},\tilde{\mathbf{w}};\mathbf{c}), pq) = 1. \tag{3.11}$$

Lemma 3.1. *Given $\varepsilon > 0$, there exists a positive number $\eta = \eta(\varepsilon, s, k)$ such that whenever $R \le P^\eta$ one has*

$$S_{s,r}(P,Q,R;\boldsymbol{\psi}) \ll \tilde{C}P^{3r-1}S_s(Q,R) + \tilde{C}Q^{3s}P^{2r+s\theta+\varepsilon} + P^{(4s-2)\theta+\varepsilon}T_{s,r}(P,Q,R,\theta;\boldsymbol{\psi}).$$

Proof. Let S_1 denote the number of solutions counted by $S_{s,r}(P,Q,R;\boldsymbol{\psi})$ such that $(\mathbf{z},\mathbf{w},\tilde{\mathbf{z}},\tilde{\mathbf{w}})$ is highly singular for $\boldsymbol{\psi}$, and let S_2 denote the number of solutions such that $(\mathbf{z},\mathbf{w},\tilde{\mathbf{z}},\tilde{\mathbf{w}})$ is not highly singular for $\boldsymbol{\psi}$, so that $S_{s,r}(P,Q,R;\boldsymbol{\psi}) = S_1 + S_2$.

(i) Suppose that $S_1 \ge S_2$, so that $S_{s,r}(P,Q,R;\boldsymbol{\psi}) \le 2S_1$. By Lemma 2.3, we see that there are $O(P^{3r-1})$ permissible choices for $\mathbf{z}, \mathbf{w}, \tilde{\mathbf{z}}$, and $\tilde{\mathbf{w}}$. Now let

$$f_{\mathbf{c}}(\boldsymbol{\alpha};Q,R) = \sum_{\substack{x,y \in \mathcal{A}(Q,R)\\ (x,y) \le P^\gamma}} e\left(\sum_{i=0}^{k} \alpha_i D_i(\mathbf{c}) x^{k-i}y^i\right).$$

For a fixed choice of $\mathbf{z}, \mathbf{w}, \tilde{\mathbf{z}}, \tilde{\mathbf{w}}, \mathbf{c}$, and $\boldsymbol{\eta}$, the number of possible choices for $\mathbf{x}, \mathbf{y}, \tilde{\mathbf{x}}$, and $\tilde{\mathbf{y}}$ is at most

$$\int_{\mathbb{T}^{k+1}} |f_{\mathbf{c}}(\boldsymbol{\alpha};Q,R)|^{2s} d\boldsymbol{\alpha} \le S_s(Q,R),$$

so we have $S_1 \ll P^{3r-1}\tilde{C}S_s(Q,R)$, which establishes the lemma in this case.

(ii) Suppose that $S_2 \ge S_1$, so that $S_{s,r}(P,Q,R;\boldsymbol{\psi}) \le 2S_2$. By rearranging variables, we see that $S_{s,r}(P,Q,R;\boldsymbol{\psi}) \ll S_3$, where S_3 denotes the number of solutions of (3.1) with (3.2), (3.3), (3.4), and (3.5), and $J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c}) \ne 0$. Then by using the Cauchy-Schwarz inequality as in the corresponding argument of Wooley [17] to manipulate the underlying mean values, we see that

$$S_{s,r}(P,Q,R;\boldsymbol{\psi}) \ll S_4,$$

where S_4 represents the number of solutions of (3.1) satisfying (3.2), (3.3), (3.4), (3.5), $J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c})\neq 0$, and $J_{2r}(\tilde{\mathbf{z}},\tilde{\mathbf{w}};\mathbf{c})\neq 0$.

We now further classify the solutions counted by S_4. Write $x\ \mathcal{D}(L)\ y$ if there is some divisor d of x with $d\leq L$ such that x/d has all of its prime divisors amongst those of y. Let S_5 denote the number of solutions counted by S_4 for which

$$x_j\ \mathcal{D}(P^\theta)\ J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c})\quad\text{or}\quad \tilde{x}_j\ \mathcal{D}(P^\theta)\ J_{2r}(\tilde{\mathbf{z}},\tilde{\mathbf{w}};\mathbf{c}) \tag{3.12}$$

or

$$y_j\ \mathcal{D}(P^\theta)\ J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c})\quad\text{or}\quad \tilde{y}_j\ \mathcal{D}(P^\theta)\ J_{2r}(\tilde{\mathbf{z}},\tilde{\mathbf{w}};\mathbf{c}) \tag{3.13}$$

for some j, and let S_6 denote the number of solutions for which neither (3.12) nor (3.13) holds for any j. Then we have

$$S_{s,r}(P,Q,R;\boldsymbol{\psi})\ll S_5+S_6,$$

and we divide into further cases.

(iii) Suppose that $S_5\geq S_6$, and further suppose that (3.12) holds. Write

$$\mathcal{S}(\mathbf{z},\mathbf{w};\mathbf{c})=\{x\in\mathcal{A}(Q,R):x\ \mathcal{D}(P^\theta)\ J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c})\},$$

and let

$$\tilde{H}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha};P,Q,R)=\sum_{\substack{\mathbf{z},\mathbf{w}\\ J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c})\neq 0}}\ \sum_{\substack{x\in\mathcal{S}(\mathbf{z},\mathbf{w};\mathbf{c})\\ y\in\mathcal{A}(Q,R)\\ (x,y)\leq P^\gamma}} e(\Xi(\boldsymbol{\alpha};x,y,\mathbf{z},\mathbf{w};\mathbf{c},\boldsymbol{\eta})),$$

where

$$\Xi(\boldsymbol{\alpha};x,y,\mathbf{z},\mathbf{w};\mathbf{c},\boldsymbol{\eta})=\sum_{i=0}^{k}\alpha_i(D_i(\mathbf{c})x^{k-i}y^i+\eta_1\psi_i(z_1,w_1;\mathbf{c})+\cdots+\eta_r\psi_i(z_r,w_r;\mathbf{c})).$$

Then

$$S_5\ll\sum_{\mathbf{c},\boldsymbol{\eta},\boldsymbol{\omega}}\int_{\mathbb{T}^{k+1}}|\tilde{H}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha};P,Q,R)\tilde{F}^*_{\mathbf{c},\boldsymbol{\omega}}(\boldsymbol{\alpha};P)f_{\mathbf{c}}(\boldsymbol{\alpha};Q,R)^{2s-1}|\,d\boldsymbol{\alpha},$$

where

$$\tilde{F}^*_{\mathbf{c},\boldsymbol{\omega}}(\boldsymbol{\alpha};P)=\sum_{\substack{\mathbf{z},\mathbf{w}\\ J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c})\neq 0}} e(\Xi(\boldsymbol{\alpha};0,0,\mathbf{z},\mathbf{w};\mathbf{c},\boldsymbol{\omega})).$$

By using the Cauchy-Schwarz inequality and considering the underlying Diophantine equations as in [17], we deduce that

$$S_{s,r}(P,Q,R;\boldsymbol{\psi})\ll\sum_{g,h,\mathbf{c}}V(g,h;\mathbf{c}),$$

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

$$\begin{aligned}\sum_{n=1}^{r}\eta_n(\psi_i(z_n,w_n;\mathbf{c})-\psi_i(\tilde{z}_n,\tilde{w}_n;\mathbf{c}))+D_i(\mathbf{c})\sum_{m=1}^{s-1}(x_m^{k-i}y_m^i-\tilde{x}_m^{k-i}\tilde{y}_m^i)&\\ =D_i(\mathbf{c})((e\tilde{x})^{k-i}\tilde{y}^i-(dx)^{k-i}y^i)\quad(0\leq i\leq k)&\end{aligned}$$

with (3.2), (3.3), (3.4), (3.5), and

$$J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c})\neq 0,\quad J_{2r}(\tilde{\mathbf{z}},\tilde{\mathbf{w}};\mathbf{c})\neq 0,\quad g|J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c}),\quad h|J_{2r}(\tilde{\mathbf{z}},\tilde{\mathbf{w}};\mathbf{c}),$$

$$1\leq d,e\leq P^{\theta},\quad x\leq Q/d,\quad \tilde{x}\leq Q/e,\quad y,\tilde{y}\leq Q,\quad s_0(x)=g,\quad s_0(\tilde{x})=h.$$

Write

$$G_{\mathbf{c},\boldsymbol{\eta},g}(\boldsymbol{\alpha};P)=\sum_{\substack{\mathbf{z},\mathbf{w}\\ g|J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c})\neq 0}} e(\Xi(\boldsymbol{\alpha};0,0,\mathbf{z},\mathbf{w};\mathbf{c},\boldsymbol{\eta}))$$

and

$$\mathcal{G}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha})=\sum_{g\leq Q}G_{\mathbf{c},\boldsymbol{\eta},g}(\boldsymbol{\alpha};P)\sum_{d\leq P^{\theta}}\sum_{\substack{x\leq Q/d\\ s_0(x)=g\\ y\leq Q}} e\left(\sum_{i=0}^{k}\alpha_i D_i(\mathbf{c})(dx)^{k-i}y^i\right).$$

Then

$$S_{s,r}(P,Q,R;\boldsymbol{\psi})\ll\sum_{\mathbf{c},\boldsymbol{\eta}}\int_{\mathbb{T}^{k+1}}|\mathcal{G}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha})^2 f_{\mathbf{c}}(\boldsymbol{\alpha};Q,R)^{2s-2}|\,d\boldsymbol{\alpha}. \tag{3.14}$$

By Cauchy's inequality, we have

$$|\mathcal{G}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha})|^2\leq\mathcal{H}_{1,\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha})\mathcal{H}_{2,\mathbf{c}}(\boldsymbol{\alpha}), \tag{3.15}$$

where

$$\mathcal{H}_{1,\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha})=\sum_{g\leq Q}|G_{\mathbf{c},\boldsymbol{\eta},g}(\boldsymbol{\alpha};P)|^2$$

and

$$\mathcal{H}_{2,\mathbf{c}}(\boldsymbol{\alpha})=\sum_{g\leq Q}\left|\sum_{d\leq P^{\theta}}\sum_{\substack{x\leq Q/d\\ s_0(x)=g}}\sum_{y\leq Q} e\left(\sum_{i=0}^{k}\alpha_i D_i(\mathbf{c})(dx)^{k-i}y^i\right)\right|^2.$$

Now by interchanging the order of summation and using Cauchy's inequality together with Lemma 2.4 as in [17], we obtain

$$\begin{aligned}\mathcal{H}_{2,\mathbf{c}}(\boldsymbol{\alpha}) &= \sum_{g\leq Q}\left|\sum_{\substack{x,y\leq Q\\ s_0(x)=g}}\sum_{\substack{d\leq P^{\theta}\\ d\leq Q/x}} e\left(\sum_{i=0}^{k}\alpha_i D_i(\mathbf{c})(dx)^{k-i}y^i\right)\right|^2\\ &\ll \sum_{g\leq Q}Q^{1+\varepsilon}\sum_{\substack{x,y\leq Q\\ s_0(x)=g}}P^{\theta}Q/x\\ &\ll Q^3P^{\theta+\varepsilon}.\end{aligned} \tag{3.16}$$

Thus an application of Hölder's inequality in (3.14) gives

$$\begin{aligned} S_{s,r} &\ll \left(\sum_{\mathbf{c},\boldsymbol{\eta}} \int_{\mathbb{T}^{k+1}} |\mathcal{H}_{1,\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha}) f_{\mathbf{c}}(\boldsymbol{\alpha})^{2s}| d\boldsymbol{\alpha}\right)^{1-\frac{1}{s}} \left(\sum_{\mathbf{c},\boldsymbol{\eta}} \int_{\mathbb{T}^{k+1}} |\mathcal{H}_{1,\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha}) \mathcal{H}_{2,\mathbf{c}}(\boldsymbol{\alpha})^{s}| d\boldsymbol{\alpha}\right)^{\frac{1}{s}} \\ &\ll Q^3 P^{\theta+\varepsilon} \left(\sum_{\mathbf{c},\boldsymbol{\eta}} \sum_{g \le Q} \int_{\mathbb{T}^{k+1}} |G_{\mathbf{c},\boldsymbol{\eta},g}(\boldsymbol{\alpha}; P)|^2 d\boldsymbol{\alpha}\right)^{\frac{1}{s}} S_{s,r}(P,Q,R;\boldsymbol{\psi})^{1-\frac{1}{s}}, \end{aligned}$$

where we have written $f_{\mathbf{c}}(\boldsymbol{\alpha})$ for $f_{\mathbf{c}}(\boldsymbol{\alpha}; Q, R)$ and used a standard estimate for the divisor function. But for a fixed choice of $\mathbf{c}, \boldsymbol{\eta}, \tilde{\mathbf{z}}$, and $\tilde{\mathbf{w}}$, the Inverse Function Theorem, in combination with Bézout's Theorem, shows that there are $O(1)$ choices of $\mathbf{z}$ and $\mathbf{w}$ satisfying

$$\sum_{n=1}^{r} \eta_n(\psi_i(z_n, w_n; \mathbf{c}) - \psi_i(\tilde{z}_n, \tilde{w}_n; \mathbf{c})) = 0 \quad (0 \le i \le k)$$

with $J_{2r}(\mathbf{z}, \mathbf{w}; \mathbf{c}) \neq 0$. Hence by another divisor estimate we see that

$$\sum_{\mathbf{c},\boldsymbol{\eta}} \sum_{g \le Q} \int_{\mathbb{T}^{k+1}} |G_{\mathbf{c},\boldsymbol{\eta},g}(\boldsymbol{\alpha}; P)|^2 d\boldsymbol{\alpha} \ll \tilde{C} P^{2r+\varepsilon},$$

and the result follows in the case where (3.12) holds. The case where (3.13) holds is handled in exactly the same manner.

(iv) Suppose that $S_6 \ge S_5$, and consider a solution counted by S_6. For a given index j, let q and p denote the largest divisors of x_j and y_j, respectively, with

$$(q, J_{2r}(\mathbf{z}, \mathbf{w}; \mathbf{c})) = (p, J_{2r}(\mathbf{z}, \mathbf{w}; \mathbf{c})) = 1.$$

Then, since neither (3.12) nor (3.13) holds, we have $q > P^\theta$ and $p > P^\theta$. Thus we can find divisors q_j of x_j and p_j of y_j such that $P^\theta < q_j, p_j \le P^\theta R$ and $(q_j p_j, J_{2r}(\mathbf{z}, \mathbf{w}; \mathbf{c})) = 1$, and we proceed similarly with the $\tilde{x}_j$ and $\tilde{y}_j$, except that we replace $J_{2r}(\mathbf{z}, \mathbf{w}; \mathbf{c})$ by $J_{2r}(\tilde{\mathbf{z}}, \tilde{\mathbf{w}}; \mathbf{c})$. Hence we see that $S_6 \ll V_1$, where V_1 denotes the number of solutions of

$$\begin{aligned} &\sum_{n=1}^{r} \eta_n(\psi_i(z_n, w_n; \mathbf{c}) - \psi_i(\tilde{z}_n, \tilde{w}_n; \mathbf{c})) \\ &\qquad + D_i(\mathbf{c}) \sum_{j=1}^{s} ((q_j u_j)^{k-i}(p_j v_j)^i - (\tilde{q}_j \tilde{u}_j)^{k-i}(\tilde{p}_j \tilde{v}_j)^i) = 0 \quad (0 \le i \le k). \end{aligned}$$

with (3.4), (3.5), and for $1 \le j \le s$

$$\begin{gathered} P^\theta < q_j, p_j, \tilde{q}_j, \tilde{p}_j \le P^\theta R, \qquad (q_j, p_j),\ (\tilde{q}_j, \tilde{p}_j) \le P^\gamma, \\ u_j, v_j, \tilde{u}_j, \tilde{v}_j \in \mathcal{A}(QP^{-\theta}, R), \qquad (u_j, v_j),\ (\tilde{u}_j, \tilde{v}_j) \le P^\gamma, \end{gathered} \tag{3.17}$$

and

$$(q_j p_j, J_{2r}(\mathbf{z}, \mathbf{w}; \mathbf{c})) = (\tilde{q}_j \tilde{p}_j, J_{2r}(\tilde{\mathbf{z}}, \tilde{\mathbf{w}}; \mathbf{c})) = 1.$$

Now write

$$F_{\mathbf{c},\boldsymbol{\eta},q}(\boldsymbol{\alpha}; P, R) = \sum_{\substack{\mathbf{z},\mathbf{w} \\ (q, J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c}))=1}} e(\Xi(\boldsymbol{\alpha}; 0, 0, \mathbf{z}, \mathbf{w}; \mathbf{c}, \boldsymbol{\eta}))$$

and

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

where

$$\mathbf{q}_j\mathbf{p}_j\boldsymbol{\alpha} = (\alpha_0 q_j^k, \alpha_1 q_j^{k-1}p_j, \dots, \alpha_k p_j^k) \quad \text{and} \quad \tilde{\mathbf{q}}_j\tilde{\mathbf{p}}_j\boldsymbol{\alpha} = (\alpha_0 \tilde{q}_j^k, \alpha_1 \tilde{q}_j^{k-1}\tilde{p}_j, \dots, \alpha_k \tilde{p}_j^k).$$

Then we have

$$V_1 \le \sum_{\mathbf{c},\boldsymbol{\eta}} \int_{\mathbb{T}^{k+1}} \sum_{\mathbf{q},\mathbf{p},\tilde{\mathbf{q}},\tilde{\mathbf{p}}} F_{\mathbf{c},\boldsymbol{\eta},\pi}(\boldsymbol{\alpha}; P, R) F_{\mathbf{c},\boldsymbol{\eta},\tilde{\pi}}(-\boldsymbol{\alpha}; P, R) \prod_{i=1}^{s} \mathcal{F}_{\mathbf{c},j}(\boldsymbol{\alpha})\, d\boldsymbol{\alpha}, \tag{3.18}$$

where

$$\pi = q_1 \cdots q_s p_1 \cdots p_s \quad \text{and} \quad \tilde{\pi} = \tilde{q}_1 \cdots \tilde{q}_s \tilde{p}_1 \cdots \tilde{p}_s,$$

and where the sum is over $\mathbf{q}$, $\mathbf{p}$, $\tilde{\mathbf{q}}$, $\tilde{\mathbf{p}}$ satisfying (3.17). Let

$$X_{\mathbf{c},\boldsymbol{\eta},j}(\boldsymbol{\alpha}) = \left| F_{\mathbf{c},\boldsymbol{\eta},\pi}(\boldsymbol{\alpha}; P, R)^2 f_{\mathbf{c}}(\mathbf{q}_j\mathbf{p}_j\boldsymbol{\alpha}; QP^{-\theta}, R)^{2s} \right|,$$

and let $Y_{\mathbf{c},\boldsymbol{\eta},j}(\boldsymbol{\alpha})$ be the analogous function for the $\tilde{q}_j$ and $\tilde{p}_j$. Then by (3.18) and two applications of Hölder's inequality (as in [17]), we obtain

$$S_6 \ll \sum_{\mathbf{q},\mathbf{p},\tilde{\mathbf{q}},\tilde{\mathbf{p}}} \prod_{j=1}^{s} \left(\sum_{\mathbf{c},\boldsymbol{\eta}} \int_{\mathbb{T}^{k+1}} X_{\mathbf{c},\boldsymbol{\eta},j}(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} \right)^{1/2s} \left(\sum_{\mathbf{c},\boldsymbol{\eta}} \int_{\mathbb{T}^{k+1}} Y_{\mathbf{c},\boldsymbol{\eta},j}(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} \right)^{1/2s}.$$

Now we observe that

$$\sum_{\mathbf{c},\boldsymbol{\eta}} \int_{\mathbb{T}^{k+1}} X_{\mathbf{c},\boldsymbol{\eta},j}(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} = W(P, Q, R, q_j, p_j),$$

where $W(P, Q, R, q, p)$ denotes the number of solutions of (3.7) with (3.4), (3.5), (3.9), (3.10), and (3.11). Thus we have

$$S_6 \ll \sum_{\mathbf{q},\mathbf{p},\tilde{\mathbf{q}},\tilde{\mathbf{p}}} \prod_{j=1}^{s} W(P, Q, R, q_j, p_j)^{1/2s}\, W(P, Q, R, \tilde{q}_j, \tilde{p}_j)^{1/2s},$$

whence by Hölder's inequality

$$\begin{aligned} S_6 &\ll \left(\sum_{\mathbf{q},\mathbf{p},\tilde{\mathbf{q}},\tilde{\mathbf{p}}} 1 \right)^{1-1/2s} \left(\sum_{\mathbf{q},\mathbf{p},\tilde{\mathbf{q}},\tilde{\mathbf{p}}} \prod_{j=1}^{s} W(P, Q, R, q_j, p_j)\, W(P, Q, R, \tilde{q}_j, \tilde{p}_j) \right)^{1/2s} \\ &\ll (P^\theta R)^{4s-2} \left(\prod_{j=1}^{2s} \sum_{\mathbf{q},\mathbf{p}} W(P, Q, R, q_j, p_j) \right)^{1/2s} \\ &\ll (P^\theta R)^{4s-2}\, T_{s,r}(P, Q, R, \theta; \boldsymbol{\psi}), \end{aligned}$$

and this completes the proof of the lemma. ☐

The following modification of Lemma 3.1 may be more useful for smaller values of k.

Lemma 3.2. *Given $\varepsilon > 0$, there exists a positive number $\eta = \eta(\varepsilon, s, k)$ such that whenever $R \le P^\eta$ one has*

$$S_{s,r}(P,Q,R;\boldsymbol{\psi}) \ll \tilde{C}P^{3r-1}S_s(Q,R) + Q^3P^{\theta+\varepsilon}\tilde{S}_{s-1,r}(P,Q,R;\boldsymbol{\psi}) \\ + P^{(4s-2)\theta+\varepsilon}T_{s,r}(P,Q,R,\theta;\boldsymbol{\psi}).$$

Proof. The only change occurs in part (iii) of the proof, where the number of solutions counted by S_5 is estimated. Substituting the bounds (3.15) and (3.16) into (3.14), we obtain

$$S_{s,r}(P,Q,R;\boldsymbol{\psi}) \ll Q^3P^{\theta+\varepsilon}\sum_{\mathbf{c},\boldsymbol{\eta}}\sum_{g\le Q}\int_{\mathbb{T}^{k+1}}|G_{\mathbf{c},\boldsymbol{\eta},g}(\boldsymbol{\alpha};P)^2 f_{\mathbf{c}}(\boldsymbol{\alpha};Q,R)^{2s-2}|\,d\boldsymbol{\alpha},$$

and the lemma follows on considering the underlying Diophantine equations and recalling a standard estimate for the divisor function. $\square$

Now let $\tilde{T}_{s,r}(P,Q,R,\theta;\boldsymbol{\psi})$ denote the number of solutions of (3.7) with (3.4), (3.5), (3.8), (3.9), (3.10) and also

$$z_n \equiv \tilde{z}_n \pmod{q^kp^k} \quad\text{and}\quad w_n \equiv \tilde{w}_n \pmod{q^kp^k} \quad (1 \le n \le r). \tag{3.19}$$

Lemma 3.3. *Given $\varepsilon > 0$, there exists a positive number $\gamma_0 = \gamma_0(\varepsilon, s, k)$ such that whenever $\gamma \le \gamma_0$ one has*

$$T_{s,r}(P,Q,R,\theta;\boldsymbol{\psi}) \ll (P^\theta R)^{2r(2r-1)+\varepsilon}\tilde{T}_{s,r}(P,Q,R,\theta;\boldsymbol{\psi}).$$

Proof. When q and p satisfy (3.8), let $\mathcal{B}_{q,p}(\mathbf{u};\mathbf{c},\boldsymbol{\eta})$ denote the set of solutions $(\mathbf{z},\mathbf{w})$ of the system of congruences

$$\Upsilon_i(\mathbf{z},\mathbf{w};\mathbf{c},\boldsymbol{\eta}) \equiv u_i \pmod{q^{k-i}p^i} \quad (0 \le i \le k) \tag{3.20}$$

with $1 \le z_n, w_n \le (qp)^k$ and $(qp, J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c})) = 1$, where

$$\Upsilon_i(\mathbf{z},\mathbf{w};\mathbf{c},\boldsymbol{\eta}) = \sum_{n=1}^r \psi_i(z_n,w_n;\mathbf{c}).$$

By Lemma 2.2 we have

$$\operatorname{card}(\mathcal{B}_{q,p}(\mathbf{u};\mathbf{c},\boldsymbol{\eta})) \ll (pq)^{r(2r-1)+\varepsilon},$$

on taking γ sufficiently small in terms of ε. Now observe that for each solution counted by $T_{s,r}(P,Q,R,\theta;\boldsymbol{\psi})$ we have

$$\Upsilon_i(\mathbf{z},\mathbf{w};\mathbf{c},\boldsymbol{\eta}) \equiv \Upsilon_i(\tilde{\mathbf{z}},\tilde{\mathbf{w}};\mathbf{c},\boldsymbol{\eta}) \pmod{q^{k-i}p^i},$$

so for each i we can classify the solutions of (3.7) according to the common residue class modulo $q^{k-i}p^i$ of $\Upsilon_i(\mathbf{z},\mathbf{w};\mathbf{c},\boldsymbol{\eta})$ and $\Upsilon_i(\tilde{\mathbf{z}},\tilde{\mathbf{w}};\mathbf{c},\boldsymbol{\eta})$. Let

$$H_{q,p}(\boldsymbol{\alpha};\mathbf{z},\mathbf{w};\mathbf{c},\boldsymbol{\eta}) = \sum_{\substack{\mathbf{x}\in[1,P]^r \\ x_n\equiv z_n(q^kp^k)}}\sum_{\substack{\mathbf{y}\in[1,P]^r \\ y_n\equiv w_n(q^kp^k)}} e\left(\sum_{i=0}^k \alpha_i\Upsilon_i(\mathbf{x},\mathbf{y};\mathbf{c},\boldsymbol{\eta})\right).$$

Then

$$T_{s,r}(P,Q,R,\theta;\boldsymbol{\psi}) \ll \sum_{q,p}\sum_{\mathbf{c},\boldsymbol{\eta}}\int_{\mathbb{T}^{k+1}}\tilde{H}_{q,p}(\boldsymbol{\alpha};\mathbf{c},\boldsymbol{\eta})|\tilde{f}_{\mathbf{c},q,p}(\boldsymbol{\alpha};QP^{-\theta},R)|^{2s}d\boldsymbol{\alpha},$$

where

$$\tilde{H}_{q,p}(\boldsymbol{\alpha};\mathbf{c},\boldsymbol{\eta}) = \sum_{u_0=1}^{q^k}\sum_{u_1=1}^{q^{k-1}p}\cdots\sum_{u_k=1}^{p^k}\left|\sum_{(\mathbf{z},\mathbf{w})\in\mathcal{B}_{q,p}(\mathbf{u};\mathbf{c},\boldsymbol{\eta})} H_{q,p}(\boldsymbol{\alpha};\mathbf{z},\mathbf{w};\mathbf{c},\boldsymbol{\eta})\right|^2$$

and

$$\tilde{f}_{\mathbf{c},q,p}(\boldsymbol{\alpha};L,R) = \sum_{\substack{x,y\in\mathcal{A}(L,R)\\(x,y)\leq P^\gamma}} e\left(\sum_{i=0}^{k}\alpha_i D_i(\mathbf{c})(qx)^{k-i}(py)^i\right).$$

Now by Cauchy's inequality,

$$\tilde{H}_{q,p}(\boldsymbol{\alpha};\mathbf{c},\boldsymbol{\eta}) \leq \sum_{u_0=1}^{q^k}\sum_{u_1=1}^{q^{k-1}p}\cdots\sum_{u_k=1}^{p^k}\mathrm{card}(\mathcal{B}_{q,p}(\mathbf{u};\mathbf{c},\boldsymbol{\eta}))\sum_{(\mathbf{z},\mathbf{w})\in\mathcal{B}_{q,p}(\mathbf{u};\mathbf{c},\boldsymbol{\eta})}|H_{q,p}(\boldsymbol{\alpha};\mathbf{z},\mathbf{w};\mathbf{c},\boldsymbol{\eta})|^2,$$

and thus

$$\begin{aligned} T_{s,r}(P,Q,R,\theta;\boldsymbol{\psi}) &\ll (P^\theta R)^{2r(2r-1)+\varepsilon}\sum_{\substack{q,p\\ \mathbf{c},\boldsymbol{\eta}}}\sum_{\substack{\mathbf{z},\mathbf{w}\\ 1\leq z_n\leq q^kp^k\\ 1\leq w_n\leq q^kp^k}}\int_{\mathbb{T}^{k+1}}|H_{q,p}|^2|\tilde{f}_{\mathbf{c},q,p}|^{2s}d\boldsymbol{\alpha}\\ &\ll (P^\theta R)^{2r(2r-1)+\varepsilon}\,\tilde{T}_{s,r}(P,Q,R,\theta;\boldsymbol{\psi}). \end{aligned}$$

This completes the proof. □

4. Efficient Differencing

Define the difference operator Δ_j^* recursively by

$$\Delta_1^*(f(x,y);h;g) = f(x+h,y+g) - f(x,y)$$

and

$$\begin{aligned}\Delta_{j+1}^*&(f(x,y);h_1,\ldots,h_{j+1};g_1,\ldots,g_{j+1})\\ &= \Delta_1^*(\Delta_j^*(f(x,y);h_1,\ldots,h_j;g_1,\ldots,g_j);h_{j+1};g_{j+1}),\end{aligned}$$

with the convention that

$$\Delta_0^*(f(x,y);\mathbf{h};\mathbf{g}) = f(x,y).$$

Further, write

$$\psi_{i,j}(z,w;\mathbf{h},\mathbf{g};\mathbf{m},\mathbf{n}) = \Delta_j^*(z^{k-i}w^i;h_1',\ldots,h_j';g_1',\ldots,g_j'),$$

where

$$h_i' = h_i(m_in_i)^k \quad\text{and}\quad g_i' = g_i(m_in_i)^k, \tag{4.1}$$

and put

$$r_j = \left[\frac{k-j+1}{2}\right]. \tag{4.2}$$

Our first task is to show that the polynomials $\psi_{i,j}$ satisfy the conditions of Lemma 2.3, so that the results of the previous section may be applied. We start by expressing Δ_j^* in terms of the more familiar difference operators Δ_j defined by

$$\Delta_1(f(x);h) = f(x+h) - f(x)$$

and

$$\Delta_{j+1}(f(x);h_1,\ldots,h_{j+1}) = \Delta_1(\Delta_j(f(x);h_1,\ldots,h_j);h_{j+1}).$$

For simplicity, we introduce the functions

$$\chi_{i,j}(z,w;\mathbf{h};\mathbf{g}) = \Delta_j^*(z^{k-i}w^i;h_1,\ldots,h_j;g_1,\ldots,g_j) \tag{4.3}$$

and observe that

$$\psi_{i,j}(z,w;\mathbf{h},\mathbf{g};\mathbf{m},\mathbf{n}) = \chi_{i,j}(z,w;\mathbf{h}',\mathbf{g}'),$$

where $\mathbf{h}'$ and $\mathbf{g}'$ are defined by (4.1). As in Section 2, we write $\mathcal{J}_d$ for the set $\{1,\ldots,d\}$, and also write $\tilde{\mathcal{A}}_d$ for the set $\mathcal{J}_d \setminus \mathcal{A}$. When $\mathcal{A} = \{i_1,\ldots,i_m\} \subset \mathcal{J}_j$ with $i_1 < \cdots < i_m$, define

$$q_m^{(i)}(w;\mathbf{g},\mathcal{A}) = \Delta_m(w^i;g_{i_1},\ldots,g_{i_m}), \tag{4.4}$$

and when $\mathcal{A}$ is as above and $\mathcal{B} = \{j_1,\ldots,j_t\} \subset \mathcal{J}_j$ with $j_1 < \cdots < j_t$, define

$$p_t^{(i)}(z;\mathbf{h},\mathcal{A},\mathcal{B}) = \Delta_t((z+h_{i_1}+\cdots+h_{i_m})^{k-i};h_{j_1},\ldots,h_{j_t}). \tag{4.5}$$

Lemma 4.1. *We have*

$$\chi_{i,j}(z,w;\mathbf{h};\mathbf{g}) = \sum_{m=0}^{j} \sum_{\substack{\mathcal{A}\subset\mathcal{J}_j \\ |\mathcal{A}|=m}} p_{j-m}^{(i)}(z;\mathbf{h},\mathcal{A},\tilde{\mathcal{A}}_j)q_m^{(i)}(w;\mathbf{g},\mathcal{A}).$$

Proof. We fix i, $\mathbf{h}$, and $\mathbf{g}$ and proceed by induction on j. For brevity, we write $\chi_{i,j}(z,w)$, $q_m(w;\mathcal{A})$, and $p_t(z;\mathcal{A},\mathcal{B})$ for the functions defined by (4.3), (4.4), and (4.5), respectively. For $j=0$ we have

$$\chi_{i,0}(z,w) = z^{k-i}w^i = p_0(z;\emptyset,\emptyset)q_0(w;\emptyset).$$

Now assume the result holds for $j-1$. Then we have

$$\chi_{i,j}(z,w) = \chi_{i,j-1}(z+h_j,w+g_j) - \chi_{i,j-1}(z,w),$$

so by the inductive hypothesis we obtain

$$\chi_{i,j}(z,w) = \sum_{m=0}^{j-1} \sum_{\substack{\mathcal{A}\subset\mathcal{J}_{j-1} \\ |\mathcal{A}|=m}} \theta_{i,j}(z,w;m;\mathcal{A}),$$

where

$$\theta_{i,j} = p_{j-1-m}(z+h_j;\mathcal{A},\tilde{\mathcal{A}}_{j-1})q_m(w+g_j;\mathcal{A}) - p_{j-1-m}(z;\mathcal{A},\tilde{\mathcal{A}}_{j-1})q_m(w;\mathcal{A}).$$

The above expression can be rewritten as

$$\theta_{i,j} = p_{j-m}(z;\mathcal{A},\tilde{\mathcal{A}}_j)q_m(w;\mathcal{A}) + p_{j-1-m}(z+h_j;\mathcal{A},\tilde{\mathcal{A}}_{j-1})q_{m+1}(w;\mathcal{A}\cup\{j\}),$$

so we have

$$\begin{aligned}
\chi_{i,j} &= \sum_{m=0}^{j-1}\Bigg(\sum_{\substack{\mathcal{A}\subset\mathcal{J}_{j-1}\\|\mathcal{A}|=m}} p_{j-m}(z;\mathcal{A};\tilde{\mathcal{A}}_j)q_m(w;\mathcal{A}) + \sum_{\substack{\mathcal{A}\subset\mathcal{J}_j\\|\mathcal{A}|=m+1\\j\in\mathcal{A}}} p_{j-(m+1)}(z;\mathcal{A};\tilde{\mathcal{A}}_j)q_{m+1}(w;\mathcal{A})\Bigg)\\
&= \sum_{m=0}^{j-1}\sum_{\substack{\mathcal{A}\subset\mathcal{J}_j\\|\mathcal{A}|=m\\j\notin\mathcal{A}}} p_{j-m}(z;\mathcal{A};\tilde{\mathcal{A}}_j)q_m(w;\mathcal{A}) + \sum_{m=1}^{j}\sum_{\substack{\mathcal{A}\subset\mathcal{J}_j\\|\mathcal{A}|=m\\j\in\mathcal{A}}} p_{j-m}(z;\mathcal{A};\tilde{\mathcal{A}}_j)q_m(w;\mathcal{A}),
\end{aligned}$$

and the lemma follows. □

Now we show that the 2×2 Jacobians satisfy the condition imposed in Lemma 2.3.

Lemma 4.2. *Suppose that $0\le j<k$ and $i_1<i_2\le k-j$. Then we have*

$$d_{i_1,i_2}(z,w;\boldsymbol{\chi}_j) = p(z)w^{i_1+i_2-1} + O_z(w^{i_1+i_2-2}),$$

where $p(z)$ is a non-trivial polynomial of degree at most $2k$.

Proof. When $i<k-j$, we have by Lemma 4.1 that

$$\frac{\partial\chi_{i,j}}{\partial z} = \frac{\partial}{\partial z}\left(\Delta_j(z^{k-i};h_1,\dots,h_j)\right)w^i + O_z(w^{i-1})$$

and

$$\frac{\partial\chi_{i,j}}{\partial w} = i\Delta_j(z^{k-i};h_1,\dots,h_j)w^{i-1} + O_z(w^{i-2}),$$

and we recall (see for example Exercise 2.1 of Vaughan [12]) that

$$\Delta_j(z^k;h_1,\dots,h_j) = k(k-1)\cdots(k-j+1)h_1\cdots h_j z^{k-j} + O(z^{k-j-1}).$$

Hence if $i_2<k-j$ then we have

$$d_{i_1,i_2}(z,w;\boldsymbol{\chi}) = p(z)w^{i_1+i_2-1} + O_z(w^{i_1+i_2-2}),$$

where the leading term of $p(z)$ is

$$\frac{(h_1\cdots h_j)^2(k-i_1)!(k-i_2)!}{(k-i_1-j)!(k-i_2-j)!}((k-i_1-j)i_2-(k-i_2-j)i_1)z^{2k-i_1-i_2-2j-1},$$

and the lemma follows in this case on noting that

$$(k-i_1-j)i_2-(k-i_2-j)i_1 = (k-j)(i_2-i_1)\ne 0.$$

Now if $i=k-j$ we obtain from Lemma 4.1 that

$$\frac{\partial\chi_{i,j}}{\partial z} = O_z(w^{i-1})$$

and

$$\frac{\partial\chi_{i,j}}{\partial w} = i(k-i)!\,h_1\cdots h_j w^{i-1} + O_z(w^{i-2}).$$

Thus if $i_2 = k - j$ then we have

$$d_{i_1,i_2} = \left(\frac{i_2(h_1 \cdots h_j)^2 (k-i_1)!(k-i_2)!}{(k-i_1-j-1)!} z^{k-i_1-j-1} + O(z^{k-i_1-j-2}) \right) w^{i_1+i_2-1} \\ + O_z(w^{i_1+i_2-2}),$$

and this completes the proof. $\square$

We now consider the effect of substituting $\psi_{i,j}(z,w;\mathbf{h},\mathbf{g};\mathbf{m},\mathbf{n})$ for $\psi_i(z,w;\mathbf{c})$ in the analysis of Section 3. For $1 \le j \le k$, suppose that $0 \le \phi_j \le 1/2k$, and put

$$M_j = P^{\phi_j}, \quad H_j = PM_j^{-2k}, \quad \text{and} \quad Q_j = P(M_1 \cdots M_j)^{-1}.$$

Further, write

$$\tilde{M}_j = \prod_{i=1}^{j} M_i \quad \text{and} \quad \tilde{H}_j = \prod_{i=1}^{j} H_i.$$

We replace (3.5) by the conditions

$$1 \le h_i, g_i \le H_i \quad (1 \le i \le j), \tag{4.6}$$

$$M_i < m_i, n_i \le M_i R, \quad \text{and} \quad (m_i, n_i) \le P^\gamma \quad (1 \le i \le j), \tag{4.7}$$

and take

$$D_i(\mathbf{m},\mathbf{n}) = \prod_{l=1}^{j} m_l^{k-i} n_l^i.$$

On replacing h_i by $h_i(m_i n_i)^k$ and g_i by $g_i(m_i n_i)^k$ in the above results, we see that $\psi_{0,j}, \ldots, \psi_{2r-1,j}$ satisfy the hypotheses of Lemma 2.3 whenever $r \le r_j$. Thus we may apply Lemma 3.1 to relate $S_{s,r_j}(P,Q_j,R;\boldsymbol{\psi}_j)$ to $T_{s,r_j}(P,Q_j,R,\phi_{j+1};\boldsymbol{\psi}_j)$. The following lemma then relates $T_{s,r_j}(P,Q_j,R,\phi_{j+1};\boldsymbol{\psi}_j)$ to $S_{s,r_{j+1}}(P,Q_j,R;\boldsymbol{\psi}_{j+1})$ and hence allows us to repeat the differencing process.

Lemma 4.3. *Suppose that $r \le 2w$ and $0 \le j < k$. Then given $\varepsilon > 0$, there exists $\eta = \eta(\varepsilon,s,k)$ such that whenever $R \le P^\eta$ one has*

$$\tilde{T}_{s,r}(P,Q_j,R,\phi_{j+1};\boldsymbol{\psi}_j) \ll P^{(3-2k\phi_{j+1})r+\varepsilon} \tilde{H}_j^2 \tilde{M}_{j+1}^2 S_s(Q_{j+1},R) \\ + P^\varepsilon H_{j+1}^{2r-2} \big(\tilde{H}_{j+1}^2 \tilde{M}_{j+1}^2 S_s(Q_{j+1},R)\big)^{1-r/2w} \big(S_{s,w}(P,Q_{j+1},R;\boldsymbol{\psi}_{j+1})\big)^{r/2w}.$$

Proof. Write $\theta = \phi_{j+1}$, and define

$$\mathcal{L}_{a,b,d}(\boldsymbol{\alpha};\mathbf{h},\mathbf{g};\mathbf{m},\mathbf{n}) = \sum_{\substack{1 \le z \le P \\ z \equiv a\,(d)}} \sum_{\substack{1 \le w \le P \\ w \equiv b\,(d)}} e\left(\sum_{i=0}^{k} \alpha_i \psi_{i,j}(z,w;\mathbf{h},\mathbf{g};\mathbf{m},\mathbf{n}) \right),$$

$$\mathcal{K}_d(\boldsymbol{\alpha};\mathbf{h},\mathbf{g};\mathbf{m},\mathbf{n}) = \sum_{a=1}^{d} \sum_{b=1}^{d} |\mathcal{L}_{a,b,d}(\boldsymbol{\alpha};\mathbf{h},\mathbf{g};\mathbf{m},\mathbf{n})|^2,$$

and

$$g_{q,p}(\boldsymbol{\alpha};\mathbf{m},\mathbf{n})=\sum_{\substack{x,y\in\mathcal{A}(Q_{j+1},R)\\(x,y)\le P^{\gamma}}} e\left(\sum_{i=0}^{k}\alpha_i D_i(\mathbf{m},\mathbf{n})(qx)^{k-i}(py)^i\right).$$

Then on considering the underlying Diophantine equations, we have

$$\tilde{T}_{s,r}\asymp\sum_{\mathbf{h},\mathbf{g},\mathbf{m},\mathbf{n}}\sum_{\substack{M_{j+1}<p,q\le M_{j+1}R\\(p,q)\le P^{\gamma}}}\int_{\mathbb{T}^{k+1}}\mathcal{K}_{q^kp^k}(\boldsymbol{\alpha};\mathbf{h},\mathbf{g};\mathbf{m},\mathbf{n})^r|g_{q,p}(\boldsymbol{\alpha};\mathbf{m},\mathbf{n})|^{2s}d\boldsymbol{\alpha}.$$

Let U_0 be the number of solutions counted by $\tilde{T}_{s,r}$ with $z_n=\tilde{z}_n$ or $w_n=\tilde{w}_n$ for some n, and let U_1 be the number of solutions in which $z_n\neq\tilde{z}_n$ and $w_n\neq\tilde{w}_n$ for all n, so that $\tilde{T}_{s,r}=U_0+U_1$.

First suppose that $U_0\ge U_1$, so that $\tilde{T}_{s,r}\ll U_0$. Then

$$U_0\ll P^{3-2k\phi_{j+1}}\sum_{\mathbf{h},\mathbf{g},\mathbf{m},\mathbf{n}}\sum_{M_{j+1}<p,q\le M_{j+1}R}\int_{\mathbb{T}^{k+1}}\mathcal{K}_{q^kp^k}(\boldsymbol{\alpha};\mathbf{h},\mathbf{g};\mathbf{m},\mathbf{n})^{r-1}|g_{q,p}(\boldsymbol{\alpha};\mathbf{m},\mathbf{n})|^{2s}d\boldsymbol{\alpha},$$

and by using Hölder's inequality twice as in [17], we find that

$$\tilde{T}_{s,r}\ll P^{(3-2k\phi_{j+1})r+\varepsilon}\tilde{H}_j^2\tilde{M}_{j+1}^2S_s(Q_{j+1},R). \tag{4.8}$$

Now suppose that $U_1\ge U_0$, so that $\tilde{T}_{s,r}\ll U_1$. Note that for each solution counted by U_1 we can write

$$\tilde{z}_n=z_n+\tilde{h}_nq^kp^k\quad\text{and}\quad\tilde{w}_n=w_n+\tilde{g}_nq^kp^k$$

for $1\le n\le r$, where $\tilde{h}_n,\tilde{g}_n$ are integers satisfying $1\le|\tilde{h}_n|,|\tilde{g}_n|\le H_{j+1}$. Thus we see that

$$U_1\le\sum_{\boldsymbol{\eta}\in\{\pm1\}^r}U_2(\boldsymbol{\eta}),$$

where $U_2(\boldsymbol{\eta})$ is the number of solutions of the system

$$\sum_{l=1}^{r}\eta_l\psi_{i,j+1}(z_l,w_l;\mathbf{h},\tilde{h}_l;\mathbf{g},\tilde{g}_l;\mathbf{m},q;\mathbf{n},p)\\ +\ D_i(\mathbf{m},\mathbf{n})q^{k-i}p^i\sum_{m=1}^{s}(u_m^{k-i}v_m^i-\tilde{u}_m^{k-i}\tilde{v}_m^i)=0\quad(0\le i\le k)$$

with $\mathbf{z},\mathbf{w},\mathbf{u},\mathbf{v},\tilde{\mathbf{u}},\tilde{\mathbf{v}},\mathbf{h},\mathbf{g},\mathbf{m},\mathbf{n}$ satisfying (3.4), (3.9), (3.10), (4.6), and (4.7), and with

$$1\le\tilde{h}_l,\tilde{g}_l\le H_{j+1}\quad(1\le l\le r),$$

$$M_{j+1}<p,q\le M_{j+1}R,\quad\text{and}\quad(q,p)\le P^{\gamma}.$$

On writing

$$G(\boldsymbol{\alpha};\tilde{h},\tilde{g};q,p)=\sum_{1\le z,w\le P}e\left(\sum_{i=0}^{k}\alpha_i\psi_{i,j+1}(z,w;\mathbf{h},\tilde{h};\mathbf{g},\tilde{g};\mathbf{m},q;\mathbf{n},p)\right),$$

we have by Hölder's inequality that

$$\begin{aligned} U_2(\boldsymbol{\eta}) &\ll \sum_{\mathbf{h},\mathbf{g},\mathbf{m},\mathbf{n}} \sum_{q,p} \int_{\mathbb{T}^{k+1}} \left| \sum_{1\le \tilde{g},\tilde{h} \le H_{j+1}} G(\boldsymbol{\alpha}; \tilde{h}, \tilde{g}, q, p) \right|^r |g_{q,p}(\boldsymbol{\alpha}; \mathbf{m}, \mathbf{n})|^{2s}\, d\boldsymbol{\alpha} \\ &\ll H_{j+1}^{2r-2} \sum_{\mathbf{h},\mathbf{g},\mathbf{m},\mathbf{n}} \sum_{q,p,\tilde{h},\tilde{g}} \int_{\mathbb{T}^{k+1}} |G(\boldsymbol{\alpha}; \tilde{h}, \tilde{g}, q, p)|^r |g_{q,p}(\boldsymbol{\alpha}; \mathbf{m}, \mathbf{n})|^{2s} d\boldsymbol{\alpha}. \end{aligned}$$

Thus on using Hölder's inequality twice more and considering the underlying Diophantine equations, we see that

$$\begin{aligned} U_2(\boldsymbol{\eta}) &\ll H_{j+1}^{2r-2} \sum_{\substack{\mathbf{h},\mathbf{g},\mathbf{m},\mathbf{n} \\ q,p,\tilde{h},\tilde{g}}} \left(\int_{\mathbb{T}^{k+1}} |G|^{2w} |g_{q,p}|^{2s} d\boldsymbol{\alpha} \right)^{r/2w} \left(\int_{\mathbb{T}^{k+1}} |g_{q,p}|^{2s} d\boldsymbol{\alpha} \right)^{1-r/2w} \\ &\ll H_{j+1}^{2r-2} \Big(S_{s,w}(P, Q_{j+1}, R; \boldsymbol{\psi}_{j+1}) \Big)^{r/2w} \Big(P^{\varepsilon} \tilde{H}_{j+1}^2 \tilde{M}_{j+1}^2 S_s(Q_{j+1}, R) \Big)^{1-r/2w}, \end{aligned}$$

and the lemma follows on combining this with (4.8). □

In analogy with Lemma 4.2 of [17], one might hope to refine the above argument to allow the factor of $P^{(3-2k\phi_{j+1})r}$ in the first term of the estimate to be replaced by P^{2r}, but it is not clear that this can be achieved. As will be seen in Section 6, such an improvement would have a significant impact on the strength of our repeated efficient differencing procedure.

5. Mean Value Estimates Based on Single Differencing

In this section, we consider estimates for $S_s(P, R)$ arising from a single efficient difference, reserving the full power of the preceding analysis for Section 6.

Suppose that $0 < \theta \le 1/2k$, write $r = r_0 = \left[\frac{k+1}{2}\right]$, and put

$$M = P^{\theta}, \quad H = PM^{-2k}, \quad \text{and} \quad Q = PM^{-1}.$$

Further, let

$$F(\boldsymbol{\alpha}; P) = \sum_{1\le z,w\le P} e(\alpha_0 z^k + \alpha_1 z^{k-1} w + \cdots + \alpha_k w^k),$$

$$G(\boldsymbol{\alpha}; q, p) = \sum_{1\le h,g\le H} \sum_{1\le z,w\le P} e\left(\sum_{i=0}^{k} \alpha_i \psi_{i,1}(z, w; h, g; q, p) \right),$$

$$g_{q,p}(\boldsymbol{\alpha}; P, Q, R) = \sum_{\substack{x,y\in\mathcal{A}(Q,R) \\ (x,y)\le P^{\gamma}}} e\left(\sum_{i=0}^{k} \alpha_i (qx)^{k-i} (py)^i \right),$$

and

$$\mathcal{M}_{s,r}(P, Q, R) = \sum_{M\le p,q\le MR} \int_{\mathbb{T}^{k+1}} \left| G(\boldsymbol{\alpha}; q, p)^r g_{q,p}(\boldsymbol{\alpha}; P, Q, R)^{2s} \right| d\boldsymbol{\alpha}.$$

We say that λ_s is a permissible exponent if for every $\varepsilon > 0$ there exists $\eta = \eta(\varepsilon, s, k)$ such that $S_s(P,R) \ll_\varepsilon P^{\lambda_s+\varepsilon}$ whenever $R \le P^\eta$. Further, we recall that the exponent Δ_s admissible if $\lambda_s = 4s - k(k+1) + \Delta_s$ is permissible.

Lemma 5.1. *Let $\theta = 1/2k$, and suppose that $s \ge k^2/(1-\theta)$. If $\Delta_s \le k(k+1)$ is an admissible exponent, then the exponent $\Delta_{s+r} = \Delta_s(1-\theta)$ is admissible.*

Proof. By Lemmata 3.1 and 3.3, we have

$$\begin{aligned} S_{s,r}(P,P,R;\boldsymbol{\psi}_0) \ll P^{3r-1}S_s(P,R) &+ P^{(3+\theta)s+2r+\varepsilon} \\ &+ P^\varepsilon M^{4s-2+2r(2r-1)}\tilde{T}_{s,r}(P,P,R,\theta;\boldsymbol{\psi}_0) \end{aligned} \tag{5.1}$$

for γ sufficiently small, and by the argument of the proof of Lemma 4.3 we have

$$\tilde{T}_{s,r}(P,P,R,\theta;\boldsymbol{\psi}_0) \ll P^{(3-2k\theta)r+\varepsilon}M^2S_s(Q,R) + \mathcal{M}_{s,r}(P,Q,R). \tag{5.2}$$

Since $\theta = 1/2k$, we have $H = 1$, so by a trivial estimate we obtain

$$\mathcal{M}_{s,r}(P,Q,R) \ll M^2P^{2r+\varepsilon}S_s(Q,R).$$

Hence on recalling Lemma 2.1 and considering the underlying Diophantine equations, we obtain from (5.1) and (5.2) that

$$\begin{aligned} S_{s+r}(P,R) \ll P^{2s+2r+\varepsilon} + S_{s,r}(P,P,R;\boldsymbol{\psi}_0) &\ll P^{3r-1}S_s(P,R) \\ &+ P^{(3+\theta)s+2r+\varepsilon} + P^{2r+\varepsilon}M^{4s+2r(2r-1)}S_s(Q,R). \end{aligned} \tag{5.3}$$

Thus, since $\lambda_s = 4s - k(k+1) + \Delta_s$ is permissible, we have

$$S_{s+r}(P,R) \ll P^{\Lambda_1+\varepsilon} + P^{\Lambda_2+\varepsilon} + P^{\Lambda_3+\varepsilon},$$

where

$$\Lambda_1 = 4(s+r) - k(k+1) - (r+1) + \Delta_s,$$

$$\Lambda_2 = 4(s+r) - k(k+1) - s(1-\theta) - 2r + k(k+1),$$

and

$$\Lambda_3 = 4(s+r) - k(k+1) + \Delta_s(1-\theta).$$

Now since $r+1 \ge \frac{k+1}{2}$ and $\Delta_s \le k(k+1)$, we have $\Delta_s\theta \le r+1$ and hence $\Lambda_1 \le \Lambda_3$. Furthermore, since $s(1-\theta) \ge k^2$ and $2r \ge k$, we have $\Lambda_2 \le \Lambda_3$. Therefore, the exponent $\Delta_{s+r} = \Delta_s(1-\theta)$ is admissible, and this completes the proof. $\square$

Proof of Theorem 1. Let s_1 be as in the statement of the theorem, and suppose that $s \ge s_1$. Choose an integer t with $s \equiv t \pmod r$ and $s_1 - r < t \le s_1$. Then since $\Delta_t = k(k+1)$ is trivially admissible, we find by repeated use of Lemma 5.1 that the exponent

$$\Delta_s = k(k+1)\left(1 - \frac{1}{2k}\right)^{(s-t)/r} \le\ k(k+1)\left(1 - \frac{1}{2k}\right)^{(s-s_1)/r}$$

is admissible, and this completes the proof.

6. Estimates Arising from Repeated Differencing

In this section, we explore the possibility of obtaining improved mean value estimates by employing our efficient differencing procedure repeatedly. As we take more differences, we must reduce the number of variables taken in a complete interval, so that the difference polynomials $\boldsymbol{\psi}_j$ will satisfy the hypotheses of Lemma 2.3. This complicates the recursion for generating admissible exponents and therefore requires some additional notation. Recall the definition of r_j from (4.2), and write

$$\Omega_j = \sum_{2r_j < l \le k+1} (k-l+1) = \tfrac{1}{2}(k-2r_j+1)(k-2r_j). \tag{6.1}$$

For convenience, we also write $r = r_0 = \left[\frac{k+1}{2}\right]$. Throughout this section, we will assume that k is taken to be sufficiently large.

Lemma 6.1. *Suppose that $u \ge k(k+1)$ and that $\Delta_u \le k(k+1)$ is an admissible exponent. For any integer j with $1 \le j \le \sqrt{k}$ and $s = u + lr$ $(l \in \mathbb{N})$, define the numbers Δ_s, θ_s, and $\phi(j,s,J)$ recursively as follows. For $l \ge 1$, set $\phi(j,s,j) = 1/2k$ and evaluate $\phi(j,s,J-1)$ for $J = j, \dots, 2$ by*

$$\phi^*(j,s,J-1) = \frac{1}{4k} + \left(\frac{1}{2} + \frac{2\Omega_{J-1} - \Delta_{s-r}}{8kr_{J-1}}\right)\phi(j,s,J) \tag{6.2}$$

and

$$\phi(j,s,J) = \min(1/2k, \phi^*(j,s,J)).$$

Finally, put

$$\Delta_s = \Delta_{s-r}(1-\theta_s) + r(2k\theta_s - 1) \tag{6.3}$$

and

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

Then Δ_s is an admissible exponent for $s = u + lr$ $(l \in \mathbb{N})$.

Proof. We start by noting that $0 < \theta_s \le 1/2k$ and that θ_s is an increasing function of s. Now let j denote the least integer with $\phi(j, s+r, 1) = \theta_{s+r}$ and write $\phi_J = \phi(j, s+r, J)$. As in the proof of [17], Theorem 6.1, we have $\phi_J < 1/2k$ whenever $J < j$. In particular, it follows that whenever $J < j$ we have $2\Omega_J - \Delta_s < 0$ and $\phi_J = \phi^*(j, s+r, J)$. We claim that $\phi_J \le \phi_{J+1}$ for $J < j$. By (6.2) and the above remarks, this is equivalent to

$$\phi_{J+1}\left(\frac{1}{2} + \frac{\Delta_s - 2\Omega_J}{8kr_J}\right) \ge \frac{1}{4k}, \tag{6.4}$$

and this is immediate when $J = j-1$, since $\Delta_s - 2\Omega_{j-1} > 0$ and $\phi_j = 1/2k$. Assuming the claim holds for J, then we see from (6.2) that

$$\phi_J\left(\frac{1}{2} + \frac{\Delta_s - 2\Omega_{J-1}}{8kr_{J-1}}\right)\left(\frac{1}{2} + \frac{\Delta_s - 2\Omega_J}{8kr_J}\right) \ge \frac{1}{4k}\left(\frac{1}{2} + \frac{\Delta_s - 2\Omega_{J-1}}{8kr_{J-1}}\right),$$

and it follows on using (6.1) that

$$\phi_J\left(\frac{1}{2} + \frac{\Delta_s - 2\Omega_{J-1}}{8kr_{J-1}}\right) \ge \frac{1}{4k}\left(\frac{r_J}{r_{J-1}}\right)\frac{2r_{J-1}(4k+1) + \Delta_s - k(k+1)}{2r_J(4k+1) + \Delta_s - k(k+1)}.$$

Since $\Delta_s \le k(k+1)$ and $r_J \le r_{J-1}$, we see that (6.4) holds with J replaced by $J-1$, and our claim follows.

For $1 \le i \le j$, we write

$$M_i = P^{\phi_i}, \quad H_i = PM_i^{-2k}, \quad \text{and} \quad Q_i = P(M_1 \cdots M_i)^{-1},$$

with the convention that $Q_0 = P$. We prove the lemma by induction on l, the case $l = 0$ having been assumed. Suppose that Δ_s is admissible, so that $S_s(Q,R) \ll Q^{\lambda_s+\varepsilon}$, where $\lambda_s = 4s - k(k+1) + \Delta_s$. We show inductively that

$$\tilde{T}_{s,r_J}(P, Q_J, R, \phi_{J+1}; \boldsymbol{\psi}_J) \ll P^{(3-2k\phi_{J+1})r_J+\varepsilon} \tilde{H}_J^2 \tilde{M}_{J+1}^2 Q_{J+1}^{\lambda_s} \tag{6.5}$$

for $J = j-1, \ldots, 0$. By Lemma 4.3 with j replaced by $j-1$, $r = r_{j-1}$ and $w = r_j$, we have that

$$\begin{aligned} \tilde{T}_{s,r_{j-1}}(P, Q_{j-1}, R, \phi_j; \boldsymbol{\psi}_{j-1}) &\ll P^{(3-2k\phi_j)r_{j-1}+\varepsilon} \tilde{H}_{j-1}^2 \tilde{M}_j^2 S_s(Q_j, R) \\ &\quad + P^{\varepsilon} H_j^{2r_{j-1}-2} (\tilde{H}_j^2 \tilde{M}_j^2 S_s(Q_j,R))^{1-\beta} (S_{s,r_j}(P, Q_j, R, \boldsymbol{\psi}_j))^{\beta}, \end{aligned}$$

where $\beta = r_{j-1}/(2r_j)$. Then on making the trivial estimate

$$S_{s,r_j}(P, Q_j, R; \boldsymbol{\psi}_j) \ll P^{4r_j+\varepsilon} \tilde{H}_j^2 \tilde{M}_j^2 S_s(Q_j, R)$$

and noting that $\phi_j = 1/2k$ and hence $H_j = 1$, we obtain

$$\begin{aligned} \tilde{T}_{s,r_{j-1}}(P, Q_{j-1}, R, \phi_j; \boldsymbol{\psi}_{j-1}) &\ll P^{2r_{j-1}+\varepsilon} \tilde{H}_{j-1}^2 \tilde{M}_j^2 S_s(Q_j, R) \\ &\ll P^{2r_{j-1}+\varepsilon} \tilde{H}_{j-1}^2 \tilde{M}_j^2 Q_j^{\lambda_s}, \end{aligned}$$

on using the outer induction hypothesis. Thus (6.5) holds in the case $J = j-1$.

Now suppose that (6.5) holds for J. Then, for γ sufficiently small, we have by Lemmata 3.1 and 3.3 that

$$S_{s,r_J}(P, Q_J, R; \boldsymbol{\psi}_J) \ll P^{\varepsilon} \tilde{H}_J^2 \tilde{M}_J^2 \left(P^{\Lambda_1} + P^{\Lambda_2} + P^{\Lambda_3}\right),$$

where

$$\Lambda_1 = 3r_J - 1 + \lambda_s(1 - \phi_1 - \cdots - \phi_J), \tag{6.6}$$

$$\Lambda_2 = 3s(1 - \phi_1 - \cdots - \phi_J) + s\phi_{J+1} + 2r_J, \tag{6.7}$$

and

$$\Lambda_3 = (4s + 2r_J(2r_J - 1))\phi_{J+1} + (3 - 2k\phi_{J+1})r_J + \lambda_s(1 - \phi_1 - \cdots - \phi_{J+1}). \tag{6.8}$$

Now since $J \le \sqrt{k}$, we have $r_J \sim k/2$, and it follows easily that $\Lambda_1 \le \Lambda_3$ and $\Lambda_2 \le \Lambda_3$ for $s \ge k(k+1)$ and k sufficiently large. Hence by Lemma 4.3 we have

$$\begin{aligned} \tilde{T}_{s,r_{J-1}}(P, Q_{J-1}, R, \phi_J; \boldsymbol{\psi}_{J-1}) &\ll P^{(3-2k\phi_J)r_{J-1}+\varepsilon} \tilde{H}_{J-1}^2 \tilde{M}_J^2 Q_J^{\lambda_s} \\ + P^{\varepsilon} H_J^{2r_{J-1}-2} &\left(\tilde{H}_J^2 \tilde{M}_J^2 Q_J^{\lambda_s}\right)^{1-\beta'} \left(P^{(3-2k\phi_{J+1})r_J+\varepsilon} M_{J+1}^{4s-2+2r_J(2r_J-1)} \tilde{H}_J^2 \tilde{M}_{J+1}^2 Q_{J+1}^{\lambda_s}\right)^{\beta'}, \end{aligned}$$

where $\beta' = r_{J-1}/(2r_J)$. The second term here is

$$\tilde{H}_{J-1}^2 \tilde{M}_J^2 Q_J^{\lambda_s} P^{\Lambda+\varepsilon},$$

where

$$\Lambda = 2r_{J-1}(1 - 2k\phi_J) + \frac{r_{J-1}}{2r_J}\big[(3 - 2k\phi_{J+1})r_J + (4s + 2r_J(2r_J - 1) - \lambda_s)\phi_{J+1}\big].$$

By (6.1) and (6.2), we have

$$(4s+2r_J(2r_J-1)-\lambda_s)\phi_{J+1}=(4kr_J+2\Omega_J-\Delta_s)\phi_{J+1}=8kr_J\phi_J-2r_J,$$

and hence

$$\Lambda=(\tfrac{5}{2}-k\phi_{J+1})r_{J-1}\le(3-2k\phi_J)r_{J-1}+r_{J-1}(k\phi_J-\tfrac{1}{2})\le(3-2k\phi_J)r_{J-1},$$

since $\phi_{J+1}\ge\phi_J$ and $\phi_J\le 1/2k$. Thus (6.5) holds with J replaced by $J-1$, so this completes the inner induction. Now we apply (6.5) with $J=0$ to obtain

$$\tilde{T}_{s,r}(P,P,R,\phi_1;\boldsymbol{\psi}_0)\ll P^{(3-2k\phi_1)r+2\phi_1+\lambda_s(1-\phi_1)+\varepsilon},$$

whence by Lemmata 2.1, 3.1, and 3.3 we have (for γ sufficiently small) that

$$S_{s+r}(P,R)\ll P^{2s+\varepsilon}+S_{s,r}(P,P,R;\boldsymbol{\psi}_0)\ll P^{\Lambda_1+\varepsilon}+P^{\Lambda_2+\varepsilon}+P^{\Lambda_3+\varepsilon},$$

where Λ_1,Λ_2, and Λ_3 are given by (6.6), (6.7), and (6.8) with $J=0$. Therefore, the exponent

$$\lambda_{s+r}=4(s+r)-k(k+1)+\Delta_s(1-\theta_{s+r})+r(2k\theta_{s+r}-1)$$

is permissible, and the desired conclusion holds with s replaced by $s+r$. This completes the proof of the lemma. ☐

Next we investigate the size of the admissible exponents supplied by Lemma 6.1.

Lemma 6.2. *Suppose that $s>k(k+1)+r$ and that Δ_{s-r} is an admissible exponent satisfying*

$$(\log k)^2<\Delta_{s-r}\le 2rk.$$

Write $\delta_{s-r}=\Delta_{s-r}/4rk$, and define δ_s to be the unique positive solution of the equation

$$\delta_s+\log\delta_s=\delta_{s-r}+\log\delta_{s-r}-\frac{3}{4k}+\frac{1}{k(\log k)^{3/2}}. \tag{6.9}$$

Then the exponent $\Delta_s=4rk\delta_s$ is admissible.

Proof. The proof is nearly identical to that of [17], Lemma 6.2. In view of (6.3), we may assume that $0\le\Delta_s\le 2rk$ and hence that $0\le\delta_s\le\frac{1}{2}$. By Lemma 6.1 with

$$j=\left[\frac{1}{2}(\log k)^{1/4}\right]+1, \tag{6.10}$$

we see that the exponent

$$\Delta_s=\Delta_{s-r}(1-\theta)+r(2k\theta-1)=4kr\delta_{s-r}-r+2rk(1-2\delta_{s-r})\theta, \tag{6.11}$$

is admissible, where $\theta=\theta_s=\phi(j,s,1)$. We note that for $1\le J<j$ one has

$$\Omega_J\le\tfrac{1}{2}J(J+1)<\tfrac{1}{2}(\log k)^{1/2},$$

so on writing ϕ_J for $\phi^*(j,s,J)$ we have

$$\phi_{J-1}\le\frac{1}{4k}+\frac{1}{2}(1-\delta')\phi_J, \tag{6.12}$$

where

$$\delta'=\frac{\Delta_{s-r}-(\log k)^{1/2}}{4kr}\ge\delta_{s-r}(1-(\log k)^{-3/2}). \tag{6.13}$$

An easy induction using (6.12) shows that

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

and therefore

$$\theta = \phi_1 \le \frac{1+2^{1-j}\delta'}{2k(1+\delta')}.$$

Write $L = (\log k)^{-3/2}$. Since the expression on the right hand side of the above inequality is a decreasing function of δ', we see from (6.10) and (6.13) that

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

for k sufficiently large. It now follows from (6.11) that

$$\frac{\Delta_s}{4rk} \le \delta_{s-r}\left(1-\frac{\frac{3}{2}-w}{2k(1+\delta_{s-r})}\right),$$

where

$$w = (1-2\delta_{s-r})(\log k)^{-3/2}.$$

Hence if δ_s is defined by (6.9), then since $\log(1-x) \le -x$ for $0 < x < 1$, we have

$$\begin{aligned}\frac{\Delta_s}{4rk} + \log\frac{\Delta_s}{4rk} &\le \delta_{s-r}\left(1-\frac{\frac{3}{2}-w}{2k(1+\delta_{s-r})}\right) + \log\delta_{s-r} - \frac{\frac{3}{2}-w}{2k(1+\delta_{s-r})}\\ &\le \delta_{s-r} + \log\delta_{s-r} - \frac{3}{4k} + \frac{1}{k(\log k)^{3/2}}\\ &= \delta_s + \log\delta_s,\end{aligned}$$

so that $\delta_s \ge \Delta_s/4rk$, since $\delta + \log\delta$ is an increasing function of δ. It follows that $4rk\delta_s$ is admissible, and this completes the proof of the lemma. □

We are now fully equipped to prove Theorem 2.

Proof of Theorem 2. We first note that the theorem is trivial when $1 \le s \le s_0$. Now when $s > s_0$, define δ_s to be the unique positive solution of the equation

$$\delta_s + \log\delta_s = 1 - \frac{3(s-s_0)}{4rk} + \frac{s-s_0}{rk(\log k)^{3/2}}. \tag{6.14}$$

We show by induction that the exponent $\Delta_s = 4kr\delta_s$ is admissible whenever $s_0 < s \le s_1$. First suppose that $s_0 < s \le s_0 + r$, and observe that the exponent

$$\Delta_s = k(k+1) \le 2r(k+1)$$

is trivially admissible. Then we have

$$\frac{\Delta_s}{4rk} \le \frac{1}{2} + \frac{1}{2k},$$

and hence

$$\frac{\Delta_s}{4rk} + \log\frac{\Delta_s}{4rk} \le \frac{3}{4} + \log\frac{3}{4} < \frac{1}{2} \le 1 - \frac{3}{4k} \le \delta_s + \log\delta_s$$

for $k \geq 2$. It it follows that the exponent $4rk\delta_s$ is admissible, since $\delta + \log \delta$ is an increasing function of δ. Now suppose that $\Delta_{s-r} = 4kr\delta_{s-r}$ is admissible, where $s_0 + r < s \leq s_1$. We have by (6.14) that $\delta_{s-r} \leq 1$ and

$$\delta_{s-r} + \log \delta_{s-r} \geq 1 - \frac{3(s_1 - s_0)}{4rk} > 1 - \log(4rk) + 2 \log \log k,$$

from which it follows that

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

Thus Lemma 6.2 shows that $\Delta_s = 4rk\gamma_s$ is admissible, where γ_s is the unique positive solution of

$$\gamma_s + \log \gamma_s = \delta_{s-r} + \log \delta_{s-r} - \frac{3}{4k} + \frac{1}{k(\log k)^{3/2}}.$$

Applying (6.14) with s replaced by $s - r$ now shows that $\gamma_s + \log \gamma_s = \delta_s + \log \delta_s$, whence $\gamma_s = \delta_s$, and the induction is complete.

The theorem now follows immediately in the case where $1 \leq s \leq s_1$, since from (6.14) and the definition of s_1 we see that

$$\log \delta_s \leq 2 - \frac{3(s - s_0)}{4rk}$$

for k sufficiently large.

Now suppose that $s > s_1$, and let U denote the largest integer with $s \equiv U \pmod r$ and $U \leq s_1$, so that $U \geq s_1 - r$. Then the exponent

$$\Delta_U = 4rke^{2-3(U-s_0)/4rk} < e^4 (\log k)^2$$

is admissible, and the theorem follows on applying Lemma 5.1 repeatedly.

We note that in the presence of the refined version of Lemma 4.3 discussed at the end of Section 4, we could replace the factor of r in the second term of (6.3) by $2r$ and the $3/4k$ term in (6.9) by $1/k$. Hence we would obtain admissible exponents that decay like $k^2 e^{-2s/k^2}$ in many cases of interest.

7. Weyl Estimates

Here we obtain the estimates for smooth Weyl sums quoted in Theorem 3 by making simple modifications in the corresponding argument of Wooley [17]. In the end, a standard application of the large sieve inequality shows that these estimates follow from the mean value estimates of Theorems 1 and 2. Let

$$\mathcal{C}_q(Q) = \{x \in \mathbb{Z} \cap [1, Q] : s_0(x) | s_0(q)\},$$

where $s_0(N)$ denotes the square-free kernel of N, write

$$\psi(x, y; \boldsymbol{\alpha}) = \sum_{i=0}^{k} \alpha_i x^{k-i} y^i, \tag{7.1}$$

and define the exponential sum

$$h_{r,v,v'}(\boldsymbol{\alpha}; L, L', R, R'; \theta, \theta') = \sum_{\substack{u \in \mathcal{A}(L,R) \\ (u,r)=1}} \sum_{\substack{u' \in \mathcal{A}(L',R') \\ (u',r)=1}} e(\psi(uv, u'v'; \boldsymbol{\alpha}) + \theta u + \theta' u').$$

Also, when π is a prime, we define a set of modified smooth numbers

$$\mathcal{B}(M, \pi, R) = \{v \in \mathbb{N} : M < v \leq M\pi,\ \pi | v,\ \text{and } p|v \Rightarrow \pi \leq p \leq R\}.$$

We have the following analogue of [17], Lemma 7.2.

Lemma 7.1. *Suppose that $\boldsymbol{\alpha} \in \mathbb{R}^{k+1}$ and $r \in \mathbb{N}$. Then, whenever*

$$R \leq M < Q \ll P \quad \textit{and} \quad R \leq M' < Q' \ll P,$$

we have

$$\sum_{\substack{x \in \mathcal{A}(Q,R) \\ y \in \mathcal{A}(Q',R) \\ (xy,r)=1}} e(\psi(x, y; \boldsymbol{\alpha})) \ll P^{\varepsilon} \max_{\substack{\pi, \pi' \leq R \\ \pi, \pi' \text{prime}}} \sup_{\theta, \theta' \in [0,1]} \sum_{\substack{v \in \mathcal{B}(M,\pi,R) \\ v' \in \mathcal{B}(M',\pi',R) \\ (vv',r)=1}} |h_{r,v,v'}(\boldsymbol{\alpha}; T, T', \pi, \pi'; \theta, \theta')| + E,$$

where $T = Q/M$, $T' = Q'/M'$, and $E \ll Q'M + QM'$.

Proof. By Lemma 10.1 of Vaughan [11], we have

$$\begin{aligned}
\sum_{\substack{x \in \mathcal{A}(Q,R) \\ y \in \mathcal{A}(Q',R) \\ (xy,r)=1}} e(\psi(x, y; \boldsymbol{\alpha})) &= \sum_{\substack{M < x \leq Q \\ x \in \mathcal{A}(Q,R) \\ (x,r)=1}} \sum_{\substack{M' < y \leq Q' \\ y \in \mathcal{A}(Q',R) \\ (y,r)=1}} e(\psi(x, y; \boldsymbol{\alpha})) + O(Q'M + QM') \\
&= \sum_{\substack{\pi, \pi' \leq R \\ \pi, \pi' \text{prime} \\ (r, \pi\pi')=1}} U(\boldsymbol{\alpha}; Q, Q', M, M', R, r, \pi, \pi') + O(Q'M + QM'),
\end{aligned}$$

where

$$\begin{aligned}
&U(\boldsymbol{\alpha}; Q, Q', M, M', R, r, \pi, \pi') \\
&\qquad = \sum_{\substack{v \in \mathcal{B}(M,\pi,R) \\ (v,r)=1}} \sum_{\substack{u \in \mathcal{A}(Q/v,\pi) \\ (u,r)=1}} \sum_{\substack{v' \in \mathcal{B}(M',\pi',R) \\ (v',r)=1}} \sum_{\substack{u' \in \mathcal{A}(Q'/v',\pi') \\ (u',r)=1}} e(\psi(uv, u'v'; \boldsymbol{\alpha})).
\end{aligned}$$

Now when $v, v' \geq M$ we can use orthogonality to write

$$\begin{aligned}
&\sum_{\substack{u \in \mathcal{A}(Q/v,\pi) \\ u' \in \mathcal{A}(Q'/v',\pi') \\ (uu',r)=1}} e(\psi(uv, u'v'; \boldsymbol{\alpha})) \\
&\qquad = \int_0^1 \int_0^1 h_{r,v,v'}(\theta, \theta') \left(\sum_{x \leq Q/v} e(-\theta x) \right) \left(\sum_{x' \leq Q'/v'} e(-\theta' x') \right) d\theta \, d\theta',
\end{aligned}$$

where we have abbreviated $h_{r,v,v'}(\boldsymbol{\alpha}; T, T', \pi, \pi'; \theta, \theta')$ by $h_{r,v,v'}(\theta, \theta')$. Thus we see that

$$U(\boldsymbol{\alpha}; Q, Q', M, M', R, r, \pi, \pi') \ll \int_0^1 \int_0^1 \sum_{\substack{v \in \mathcal{B}(M,\pi,R) \\ v' \in \mathcal{B}(M',\pi',R) \\ (vv',r)=1}} |h_{r,v,v'}(\theta, \theta')| \min(Q/M, \|\theta\|^{-1}) \min(Q'/M', \|\theta'\|^{-1})\, d\theta\, d\theta',$$

and the lemma follows on noting that

$$\int_0^1 \min(X, \|\theta\|^{-1}) d\theta \ll 1 + \log X$$

for $X \geq 1$. □

Theorem 3 is an easy consequence of the following lemma.

Lemma 7.2. *Suppose that $0 < \lambda \leq \frac{1}{2}$, and write $M = P^\lambda$. Let j be an integer with $0 \leq j \leq k$, and let $\boldsymbol{\alpha} \in \mathbb{R}^{k+1}$. Suppose that $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $(a, q) = 1$, $|q\alpha_j - a| \leq \frac{1}{2}(MR)^{-k}$, $q \leq 2(MR)^k$, and either $|q\alpha_j - a| > MP^{-k}$ or $q > MR$. Then whenever s is a natural number with $2s > \max(j, k-j)$ and the exponent Δ_s is admissible we have*

$$f(\boldsymbol{\alpha}; P, R)^{2s} \ll P^{4s+\varepsilon} M^{-1} (P/M)^{\Delta_s}.$$

Proof. By Lemma 2.4, along with a standard estimate for the divisor function, we see that $\operatorname{card}(\mathcal{C}_q(X)) \ll X^\varepsilon$ whenever $\log q \ll \log X$, and it follows that

$$\begin{aligned} f(\boldsymbol{\alpha}; P, R) &= \sum_{d,e \in \mathcal{C}_q(P) \cap \mathcal{A}(P,R)} \sum_{\substack{x \in \mathcal{A}(P/d,R) \\ y \in \mathcal{A}(P/e,R) \\ (xy,q)=1}} e(\psi(xd, ye; \boldsymbol{\alpha})) \\ &\ll P^\varepsilon \max_{d,e \in \mathcal{C}_q(M/R)} \left| \sum_{\substack{x \in \mathcal{A}(P/d,R) \\ y \in \mathcal{A}(P/e,R) \\ (xy,q)=1}} e(\psi(xd, ye; \boldsymbol{\alpha})) \right| + P^{1+\varepsilon}(PR/M). \end{aligned}$$

Thus by Lemma 7.1 there exist $d, e \in \mathcal{C}_q(M/R)$, $\theta, \theta' \in [0, 1]$ and primes $\pi, \pi' \leq R$ such that

$$f(\boldsymbol{\alpha}; P, R) \ll P^{2+\varepsilon} M^{-1} + P^\varepsilon g(\boldsymbol{\alpha}; d, e, \pi, \pi', \theta, \theta'), \tag{7.2}$$

where

$$g(\boldsymbol{\alpha}; d, e, \pi, \pi', \theta, \theta') = \sum_{\substack{v \in \mathcal{B}(M/d,\pi,R) \\ (v,q)=1}} \sum_{\substack{v' \in \mathcal{B}(M/e,\pi',R) \\ (v',q)=1}} |h_{r,vd,v'e}(\boldsymbol{\alpha}; P/M, P/M, \pi, \pi'; \theta, \theta')|.$$

Let $J(q, v, d, e, h)$ denote the number of solutions of the congruence $(vd)^{k-j}(xe)^j \equiv h \pmod{q}$ with $1 \leq x \leq q$ and $(x, q) = 1$. When $(v, q) = 1$, a solution x counted by $J(q, v, d, e, h)$ satisfies $d^{k-j} e^j x^j \equiv h' \pmod{q}$, and we then necessarily have $(h', q) | d^{k-j} e^j$. In this instance, a simple application of the Chinese Remainder Theorem shows that

$$J(q, v, d, e, h) \ll q^\varepsilon d^{k-j} e^j.$$

Thus for any fixed v with $(v,q)=1$, we may divide the integers v' with $M/e<v'\le MR/e$ and $(v',q)=1$ into $L\ll q^\varepsilon d^{k-j}e^j$ classes $\mathcal{V}_1,\dots,\mathcal{V}_L$ such that, whenever $v_1',v_2'\in\mathcal{V}_r$ and $(vd)^{k-j}(v_1'e)^j\equiv(vd)^{k-j}(v_2'e)^j \pmod q$, we have $v_1'\equiv v_2' \pmod q$.

Now put $Q=P/M$, and write $c_{\mathbf{y}}$ for the number of solutions of the system

$$\sum_{i=1}^s u_i^{k-j}(u_i')^j=y_j \quad (0\le j\le k)$$

with

$$u_i\in\mathcal{A}(Q,\pi)\quad\text{and}\quad u_i'\in\mathcal{A}(Q,\pi')\quad(1\le i\le s)$$

and

$$(u_i,q)=(u_i',q)=1\qquad(1\le i\le s).$$

Further, write $g(\boldsymbol{\alpha})$ for $g(\boldsymbol{\alpha};d,e,\pi,\pi',\theta,\theta')$. Then for some r with $1\le r\le L$ we have by Hölder's inequality that

$$|g(\boldsymbol{\alpha})|^{2s}\ll P^\varepsilon d^{k-j}e^j(M^2R^2/de)^{2s-1}\sum_{\substack{v\in\mathcal{B}(M/d,\pi,R)\\(v,q)=1}}\sum_{v'\in\mathcal{V}_r}\left|\sum_{\mathbf{y}}b_{\mathbf{y}}e(\psi(vd,v'e;\boldsymbol{\alpha}\mathbf{y}))\right|^2,$$

where $|b_{\mathbf{y}}|\le c_{\mathbf{y}}$. Here we have written $\boldsymbol{\alpha}\mathbf{y}=(\alpha_0y_0,\dots,\alpha_ky_k)$, and the summation is over $\mathbf{y}$ with $1\le y_i\le sQ^k$. Applying Cauchy's inequality, we obtain

$$|g(\boldsymbol{\alpha})|^{2s}\ll P^\varepsilon M^{4s-2}Q^{k^2}\sum_{\mathbf{y}}{}^{*}\sum_{\substack{v\in\mathcal{B}(M/d,\pi,R)\\(v,q)=1}}\sum_{v'\in\mathcal{V}_r}\left|\sum_{y_j}b_{\mathbf{y}}e(\alpha_j(vd)^{k-j}(v'e)^jy_j)\right|^2, \tag{7.3}$$

where $\sum^*$ denotes the sum over y_i with $i\ne j$.

We now show that the quantities $\alpha_j(vd)^{k-j}(v'e)^j$ are well-spaced modulo 1 as v' runs through the set $\mathcal{V}_r$, and it is here that we use the "minor arc" conditions on α_j imposed in the statement of the lemma. Fix $v\in\mathcal{B}(M/d,\pi,R)$, and note that if $v_1',v_2'\in\mathcal{V}_r$ and $v_1'\not\equiv v_2'$ $\pmod q$ then since $|q\alpha_j-a|\le\frac12(MR)^{-k}$ we have

$$\begin{aligned}\left\|\alpha_j((vd)^{k-j}(v_1'e)^j-(vd)^{k-j}(v_2'e)^j)\right\| &\ge \left\|\frac{a}{q}\left((vd)^{k-j}(v_1'e)^j-(vd)^{k-j}(v_2'e)^j\right)\right\|-\frac{1}{2q}\\ &\ge \frac{1}{2q}.\end{aligned}$$

In particular, if $q>MR/e$, then the elements of $\mathcal{V}_r$ are distinct modulo q, so the $\alpha_j(vd)^{k-j}(v'e)^j$ with $v'\in\mathcal{V}_r$ are spaced at least $\frac12q^{-1}$ apart. Thus it suffices to consider the case when v_1' and v_2' are distinct elements of $\mathcal{V}_r$ with $v_1'\equiv v_2' \pmod q$ and $q\le MR/e$. In this case we have

$$\begin{aligned}\left\|\alpha_j((vd)^{k-j}(v_1'e)^j-(vd)^{k-j}(v_2'e)^j)\right\| &= \left\|\left(\alpha_j-\frac{a}{q}\right)(vd)^{k-j}e^j((v_1')^j-(v_2')^j)\right\|\\ &= \left|\alpha_j-\frac{a}{q}\right|(vd)^{k-j}e^j\left|(v_1')^j-(v_2')^j\right|.\end{aligned}$$

Now since $|q\alpha_j - a| > MP^{-k}$ and $v_1' - v_2'$ is a nonzero multiple of q, we get

$$\left\|\alpha_j((vd)^{k-j}(v_1'e)^j - (vd)^{k-j}(v_2'e)^j)\right\| \geq MP^{-k}(vd)^{k-j}e^j(v_1')^{j-1} \geq (P/M)^{-k},$$

and thus on applying the large sieve inequality to (7.3) we obtain

$$g(\boldsymbol{\alpha}; d, e, \pi, \pi', \theta, \theta')^{2s} \ll P^\varepsilon M^{4s-2}(P/M)^{k^2}(q + (P/M)^k) \sum_{v \in \mathcal{B}(M/d,\pi,R)} \sum_{\mathbf{y}} |b_{\mathbf{y}}|^2.$$

But $\sum_{\mathbf{y}} |b_{\mathbf{y}}|^2 \leq S_s(P/M, R)$ and $q \leq 2(MR)^k \ll (P/M)^k$ so on recalling (7.2) we have

$$\begin{aligned} f(\boldsymbol{\alpha}; P, R)^{2s} &\ll P^{4s+\varepsilon}M^{-2s} + P^\varepsilon M^{4s-1}(P/M)^{k^2}(P/M)^k(P/M)^{4s-k(k+1)+\Delta_s} \\ &\ll P^{4s+\varepsilon}M^{-1}(P/M)^{\Delta_s}, \end{aligned}$$

as required. □

Proof of Theorem 3. Suppose that $\boldsymbol{\alpha} \in \mathfrak{m}_{\lambda(k+1)}$ and write $M = P^\lambda$. By Dirichlet's Theorem there exist $b_i \in \mathbb{Z}$ and $q_i \in \mathbb{N}$ with $(b_i, q_i) = 1$ such that

$$|q_i\alpha_i - b_i| \leq \frac{1}{2}(MR)^{-k} \quad \text{and} \quad q_i \leq 2(MR)^k \quad (0 \leq i \leq k).$$

If for some j we have either

$$|\alpha_j - b_j/q_j| > q_j^{-1}MP^{-k} \quad \text{or} \quad q_j > MR,$$

then the desired conclusion follows from Lemma 7.2. Otherwise, write $q = [q_0, \ldots, q_k]$ and $a_i = b_iq/q_i$. Then $(a_0, \ldots, a_k, q) = 1$, and for each i we have

$$q \leq q_i(MR)^k \leq (MR)^{k+1} = P^{\lambda(k+1)}R^{k+1}$$

and

$$|\alpha_i - a_i/q| \leq q^{-1}(MR)^kMP^{-k} = q^{-1}P^{\lambda(k+1)-k}R^k.$$

This contradicts the assumption that $\boldsymbol{\alpha} \in \mathfrak{m}_{\lambda(k+1)}$ and hence completes the proof.

Proof of Corollary 3.1. We apply Theorem 3 with $\lambda = \frac{1}{2(k+1)}$. By (1.5), we have

$$\sigma(\lambda) = \max_{2s \geq k+1} \frac{1 - (2k+1)\Delta_s}{4s(k+1)}.$$

Then on taking

$$s = \left[\left(\frac{7}{3}\log 4rk + 2\log\log k + 8\right)rk\right] + 1 \sim \frac{7}{3}k^2\log k,$$

we have by Theorem 2 that the exponent

$$\Delta_s = e^4(\log k)^2e^{-(s-s_1)/2rk} \leq \frac{1}{k(\log k)^{1/3}}$$

is admissible. It follows that

$$\sigma(\lambda) \geq \frac{1 + O((\log k)^{-1/3})}{\frac{28}{3}k^3(\log k + O(\log\log k))} \sim \left(\frac{28}{3}k^3\log k\right)^{-1}.$$

We remark that the proof of Lemma 7.2, with trivial modifications, may be applied to more general exponential sums of the form

$$f(\boldsymbol{\alpha}; P, Q, R) \sum_{x\in\mathcal{A}(P,R)} \sum_{y\in\mathcal{A}(Q,R)} e(\alpha_0 x^k + \alpha_1 x^{k-1} y + \cdots + \alpha_k y^k),$$

provided that $P \asymp Q$, and hence Theorem 3 and Corollary 3.1 hold in this case as well. This observation will be useful in the analysis of Section 10.

8. Generating Function Asymptotics

In this section, we derive the asymptotic formulas for our generating functions, which will be required to handle the major arcs in our subsequent applications of the circle method.

As is now familiar in the applications of smooth numbers to additive number theory, one can only obtain asymptotics for the exponential sum $f(\boldsymbol{\alpha}; P, R)$ on a very thin set of major arcs, so it is necessary to introduce sums over a complete interval in order to facilitate a pruning procedure. Thus we write

$$F(\boldsymbol{\alpha}) = \sum_{1\le x,y\le P} e(\alpha_0 x^k + \alpha_1 x^{k-1} y + \cdots + \alpha_k y^k),$$

and we also define

$$S(q, \mathbf{a}) = \sum_{1\le x,y\le q} e\left(\frac{a_0 x^k + a_1 x^{k-1} y + \cdots + a_k y^k}{q}\right),$$

$$v(\boldsymbol{\beta}) = \int_0^P \int_0^P e(\beta_0 \gamma^k + \beta_1 \gamma^{k-1} \nu + \cdots + \beta_k \nu^k)\, d\gamma\, d\nu, \tag{8.1}$$

and

$$V(\boldsymbol{\alpha}) = V(\boldsymbol{\alpha}; q, \mathbf{a}) = q^{-2} S(q, \mathbf{a}) v(\boldsymbol{\alpha} - \mathbf{a}/q).$$

Lemma 8.1. *When $\alpha_i = a_i/q + \beta_i$ for $0 \le i \le k$, one has*

$$F(\boldsymbol{\alpha}) - V(\boldsymbol{\alpha}) \ll q^2 + qP^{k+1}\left(|\beta_0| + \cdots + |\beta_k|\right).$$

Proof. On sorting the terms into arithmetic progressions modulo q, we have

$$F(\boldsymbol{\alpha}) = \sum_{r=1}^{q} \sum_{s=1}^{q} e\left(\frac{a_0 r^k + \cdots + a_k s^k}{q}\right) \sum_{0\le i\le \frac{P-r}{q}} \sum_{0\le j\le \frac{P-s}{q}} e(\psi(iq + r, jq + s; \boldsymbol{\beta})),$$

where $\psi(x, y; \boldsymbol{\alpha})$ is as in (7.1). Thus on making the change of variables $\gamma = qz + r$ and $\nu = qw + s$ in (8.1), we obtain

$$F(\boldsymbol{\alpha}) - V(\boldsymbol{\alpha}) = \sum_{1\le r,s\le q} e\left(\frac{a_0 r^k + \cdots + a_k s^k}{q}\right) \left\{ \sum_{i,j} \int_i^{i+1} \int_j^{j+1} H(z, w)\, dz\, dw + O(1) \right\},$$

where

$$H(z, w) = H(z, w; r, s; i, j; \boldsymbol{\beta}) = e(\psi(iq + r, jq + s; \boldsymbol{\beta})) - e(\psi(qz + r, qw + s; \boldsymbol{\beta})).$$

Using the mean value theorem, we find that

$$H(z, w) \ll qP^{k-1}\left(|\beta_0| + \cdots + |\beta_k|\right)$$

when $(z,w) \in [i,i+1] \times [j,j+1]$, and hence

$$F(\boldsymbol{\alpha}) - V(\boldsymbol{\alpha}) \ll q^2(1 + q^{-1}P^{k+1}(|\beta_0| + \cdots + |\beta_k|)),$$

from which the lemma follows. □

We now begin to analyze the sum $f(\boldsymbol{\alpha}; P, R)$. First we record a partial summation lemma analogous to Lemma 2.6 of Vaughan [12].

Lemma 8.2. *Let $c_{m,n}$ be arbitrary complex numbers, and suppose that $F(x,y)$ has continuous partial derivatives on $[0,X] \times [0,Y]$. Then*

$$\sum_{\substack{m \le X \\ n \le Y}} c_{m,n} F(m,n) = \sum_{\substack{m \le X \\ n \le Y}} c_{m,n} \big(F(X,n) + F(m,Y) - F(X,Y)\big) + \int_0^X \int_0^Y \frac{\partial^2}{\partial\gamma\partial\nu} F(\gamma,\nu) \left(\sum_{\substack{m \le \gamma \\ n \le \nu}} c_{m,n} \right) d\nu\, d\gamma.$$

Proof. Write $F_\gamma(\nu) = \frac{\partial}{\partial\gamma} F(\gamma,\nu)$. Then we have

$$F_\gamma(n) = F_\gamma(Y) - \int_n^Y \frac{\partial}{\partial\nu} F_\gamma(\nu)\, d\nu$$

and

$$F(m,n) = F(X,n) - \int_m^X F_\gamma(n)\, d\gamma.$$

Thus we can write

$$F(m,n) = F(X,n) - \int_m^X F_\gamma(Y)\, d\gamma + \int_m^X \int_n^Y \frac{\partial^2}{\partial\gamma\partial\nu} F(\gamma,\nu)\, d\nu\, d\gamma,$$

and the lemma follows on summing over m and n and interchanging the order of integration and summation in the last term. □

Using the well-known asymptotics for card$(\mathcal{A}(X,R))$ in terms of Dickman's ρ function, we can record the following lemma.

Lemma 8.3. *Let τ be a fixed number, and suppose that $R \le m, n \le R^\tau$. Then*

$$\sum_{\substack{x \in \mathcal{A}(m,R) \\ y \in \mathcal{A}(n,R)}} 1 = \rho\left(\frac{\log m}{\log R}\right) \rho\left(\frac{\log n}{\log R}\right) mn + O\left(\frac{mn}{\log R}\right).$$

Proof. By Lemma 5.3 of Vaughan [11], we have

$$\sum_{x \in \mathcal{A}(X,R)} 1 = \rho\left(\frac{\log X}{\log R}\right) X + O\left(\frac{X}{\log X}\right)$$

whenever $R \le X \le R^\tau$, and the result follows immediately. □

Now let W be a parameter at our disposal, and write

$$\mathfrak{N}(q,\mathbf{a}) = \{\boldsymbol{\alpha} \in \mathbb{T}^{k+1} : |\alpha_i - a_i/q| \le WP^{-k} \ (0 \le i \le k)\} \tag{8.2}$$

whenever $q \le W$ and $(q, a_0, \ldots, a_k) = 1$. Further, let $R = P^\eta$, and write

$$w(\boldsymbol{\beta}) = \int_R^P \int_R^P \rho\left(\frac{\log \gamma}{\log R}\right) \rho\left(\frac{\log \nu}{\log R}\right) e(\beta_0 \gamma^k + \cdots + \beta_k \nu^k)\, d\gamma\, d\nu. \tag{8.3}$$

Lemma 8.4. *Suppose that $\boldsymbol{\alpha} \in \mathfrak{N}(q,\mathbf{a})$ with $q \le R$, and write $\beta_i = \alpha_i - a_i/q$. Then we have*

$$f(\boldsymbol{\alpha}; P, R) = q^{-2} S(q,\mathbf{a}) w(\boldsymbol{\beta}) + O\left(\frac{q^2 P^2 W^2}{\log P}\right).$$

Proof. By arguing as in the proof of Vaughan [11], Lemma 5.4, we obtain

$$\sum_{\substack{x \in \mathcal{A}(m,R) \\ x \equiv r\ (q)}} \sum_{\substack{y \in \mathcal{A}(n,R) \\ y \equiv s\ (q)}} 1 = \frac{1}{q^2} \sum_{\substack{x \in \mathcal{A}(m,R) \\ y \in \mathcal{A}(n,R)}} 1 + O\left(\frac{P^2}{\log P}\right)$$

whenever $R \le m, n \le P$, and hence by Lemma 8.3 we have

$$\begin{aligned} \sum_{\substack{x \in \mathcal{A}(m,R) \\ y \in \mathcal{A}(n,R)}} e\left(\frac{a_0 x^k + \cdots + a_k y^k}{q}\right) &= q^{-2} S(q,\mathbf{a}) \sum_{\substack{x \in \mathcal{A}(m,R) \\ y \in \mathcal{A}(n,R)}} 1 + O\left(\frac{q^2 P^2}{\log P}\right) \\ &= q^{-2} S(q,\mathbf{a}) \rho\left(\frac{\log m}{\log R}\right) \rho\left(\frac{\log n}{\log R}\right) mn + E_1, \end{aligned}$$

where $E_1 \ll q^2 P^2 / \log P$. Now let $\mathcal{B} = \mathcal{A}(P,R) \times \mathcal{A}(P,R)$, and write $1_{\mathcal{B}}$ for the characteristic function of $\mathcal{B}$. Then by taking

$$c_{x,y} = e\left(\frac{a_0 x^k + \cdots + a_k y^k}{q}\right) 1_{\mathcal{B}}(x,y) \quad \text{and} \quad F(x,y) = e(\beta_0 x^k + \cdots + \beta_k y^k)$$

in Lemma 8.2 we find that

$$f(\boldsymbol{\alpha}; P, R) = \sum_{1 \le x,y \le P} c_{x,y} F(x,y) = S_0 - S_1 + S_2, \tag{8.4}$$

where

$$S_0 = \sum_{x,y \in \mathcal{A}(P,R)} e\left(\frac{a_0 x^k + \cdots + a_k y^k}{q}\right) \big(e(\beta_0 P^k + \cdots + \beta_k y^k) + e(\beta_0 x^k + \cdots + \beta_k P^k)\big),$$

$$S_1 = \sum_{x,y \in \mathcal{A}(P,R)} e\left(\frac{a_0 x^k + \cdots + a_k y^k}{q}\right) e(\beta_0 P^k + \cdots + \beta_k P^k),$$

and

$$S_2 = \int_0^P \int_0^P \frac{\partial^2}{\partial \gamma \partial \nu} \big(e(\beta_0 \gamma^k + \cdots + \beta_k \nu^k)\big) \sum_{\substack{x \in \mathcal{A}(\gamma,R) \\ y \in \mathcal{A}(\nu,R)}} e\left(\frac{a_0 x^k + \cdots + a_k y^k}{q}\right) d\nu\, d\gamma.$$

From our observations above, we see immediately that

$$S_1 = q^{-2} S(q,\mathbf{a}) P^2 \rho(1/\eta)^2 e(\beta_0 P^k + \cdots + \beta_k P^k) + O\left(\frac{q^2 P^2}{\log P}\right). \tag{8.5}$$

We next observe that, by equation (8.13) of Wooley [13], one has

$$\sum_{x \in \mathcal{A}(m,R)} e\left(\frac{a_0 x^k + \cdots + a_k y^k}{q}\right) = q^{-1} S(q,\mathbf{a};y)\, m\, \rho\left(\frac{\log m}{\log R}\right) + O\left(\frac{qP}{\log P}\right), \tag{8.6}$$

where

$$S(q,\mathbf{a};y) = \sum_{1 \le x \le q} e\left(\frac{a_0 x^k + a_1 x^{k-1} y + \cdots + a_k y^k}{q}\right).$$

If we write $S_0 = S_3 + S_4$, then by (8.6) we have

$$\begin{aligned} S_3 &= \sum_{x,y \in \mathcal{A}(P,R)} e\left(\frac{a_0 x^k + \cdots + a_k y^k}{q}\right) e(\beta_0 P^k + \cdots + \beta_k y^k) \\ &= q^{-1} \rho(1/\eta) P \sum_{y \in \mathcal{A}(P,R)} S(q,\mathbf{a};y) e(\beta_0 P^k + \cdots + \beta_k y^k) + O\left(\frac{qP^2}{\log P}\right), \end{aligned}$$

and then by partial summation

$$\begin{aligned} S_3 = {} & q^{-1} P\, \rho(1/\eta)\, T(P)\, e(\beta_0 P^k + \cdots + \beta_k P^k) \\ & - q^{-1} P\, \rho(1/\eta) \int_R^P T(\nu) \frac{\partial}{\partial \nu} \left(e(\beta_0 P^k + \cdots + \beta_k \nu^k) \right) d\nu + O\left(\frac{qP^2}{\log P}\right), \end{aligned}$$

where

$$T(\nu) = \sum_{y \in \mathcal{A}(\nu,R)} S(q,\mathbf{a};y).$$

But on using the obvious analogue of (8.6) we find that

$$T(\nu) = q^{-1} S(q,\mathbf{a})\, \nu \rho\left(\frac{\log \nu}{\log R}\right) + O\left(\frac{q^2 P}{\log P}\right),$$

and since $\boldsymbol{\alpha} \in \mathfrak{N}(q,\mathbf{a})$ we have

$$\frac{\partial}{\partial \nu} \left(e(\beta_0 P^k + \cdots + \beta_k \nu^k) \right) \ll W/P.$$

Therefore we obtain

$$S_3 = Q\, \rho(1/\eta)\, P\, e(\beta_0 P^k + \cdots + \beta_k P^k) - Q\, I(P) + O\left(\frac{qP^2 W}{\log P}\right), \tag{8.7}$$

where $Q = q^{-2} S(q,\mathbf{a}) \rho(1/\eta) P$ and

$$I(\gamma) = \int_R^P \nu\, \rho\left(\frac{\log \nu}{\log R}\right) \frac{\partial}{\partial \nu} \left(e(\beta_0 \gamma^k + \cdots + \beta_k \nu^k) \right)\, d\nu.$$

Integration by parts yields

$$\begin{aligned} I(\gamma) &= \rho(1/\eta)P\, e(\beta_0\gamma^k + \cdots + \beta_k P^k) \\ &\qquad - \int_R^P e(\beta_0\gamma^k + \cdots + \beta_k\nu^k)\, \frac{\partial}{\partial\nu}\left(\nu\rho\left(\frac{\log\nu}{\log R}\right)\right) d\nu + O(R), \end{aligned}$$

but

$$\frac{\partial}{\partial\nu}\left(\nu\rho\left(\frac{\log\nu}{\log R}\right)\right) = \rho\left(\frac{\log\nu}{\log R}\right) + \frac{1}{\log R}\,\rho'\left(\frac{\log\nu}{\log R}\right) = \rho\left(\frac{\log\nu}{\log R}\right) + O\left(\frac{1}{\log P}\right),$$

since $\rho'(x) \ll 1$. Thus we have

$$\begin{aligned} I(\gamma) &= \rho(1/\eta)P\, e(\beta_0\gamma^k + \cdots + \beta_k P^k) \\ &\qquad - \int_R^P e(\beta_0\gamma^k + \cdots + \beta_k\nu^k)\, \rho\left(\frac{\log\nu}{\log R}\right) d\nu + E_2(\gamma), \end{aligned} \tag{8.8}$$

where $E_2(\gamma) \ll P/\log P$, so it follows from (8.7) that

$$S_3 = Q\int_R^P \rho\left(\frac{\log\nu}{\log R}\right) e(\beta_0 P^k + \cdots + \beta_k\nu^k)\, d\nu + O\left(\frac{qP^2W}{\log P}\right). \tag{8.9}$$

Moreover, an identical argument shows that

$$S_4 = Q\int_R^P \rho\left(\frac{\log\gamma}{\log R}\right) e(\beta_0\gamma^k + \cdots + \beta_k P^k)\, d\gamma + O\left(\frac{qP^2W}{\log P}\right). \tag{8.10}$$

We now deal with S_2. A simple calculation shows that

$$\frac{\partial^2}{\partial\gamma\partial\nu}\left(e(\beta_0\gamma^k + \cdots + \beta_k\nu^k)\right) \ll W^2/P^2.$$

when $|\beta_i| \le WP^{-k}$, and it follows easily from the calculation at the beginning of the proof that

$$\begin{aligned} S_2 = \int_R^P\int_R^P \frac{\partial^2}{\partial\gamma\partial\nu}\left(e(\beta_0\gamma^k + \cdots + \beta_k\nu^k)\right) q^{-2}S(q,\mathbf{a})\, \rho\left(\frac{\log\gamma}{\log R}\right)\rho\left(\frac{\log\nu}{\log R}\right) \gamma\nu\, d\gamma\, d\nu \\ + O\left(\frac{q^2P^2W^2}{\log P}\right). \end{aligned}$$

After interchanging the order of differentiation and integration, we can write

$$S_2 = q^{-2}S(q,\mathbf{a})\int_R^P \gamma\,\rho\left(\frac{\log\gamma}{\log R}\right) I'(\gamma)\, d\gamma + O\left(\frac{q^2P^2W^2}{\log P}\right),$$

and on integrating by parts we get

$$S_2 = q^{-2}S(q,\mathbf{a})\left(P\,\rho(1/\eta)\, I(P) - \int_R^P I(\gamma)\,\rho\left(\frac{\log\gamma}{\log R}\right) d\gamma\right) + O\left(\frac{q^2P^2W^2}{\log P}\right).$$

Then from (8.8) we finally obtain

$$S_2 = q^{-2}S(q,\mathbf{a})w(\boldsymbol{\beta}) + E_3,$$

where

$$\begin{aligned}E_3 = q^{-2}S(q,\mathbf{a})\rho(1/\eta)^2P^2e(\beta_0P^k+\cdots+\beta_kP^k)\\ - Q\int_R^P \rho\left(\frac{\log\nu}{\log R}\right)e(\beta_0P^k+\cdots+\beta_k\nu^k)\,d\nu\\ - Q\int_R^P \rho\left(\frac{\log\gamma}{\log R}\right)e(\beta_0\gamma^k+\cdots+\beta_kP^k)\,d\gamma+O\left(\frac{q^2P^2W^2}{\log P}\right),\end{aligned}$$

and the lemma follows on recalling (8.4), (8.5), (8.9), and (8.10). □

9. A Multidimensional Analogue of Waring's Problem

Here we establish Theorem 4 by a fairly straightforward application of the Hardy-Littlewood method. Let P be a large positive number, and put $R=P^\eta$, where $\eta\le\eta_0(\varepsilon,k)$. Let $F(\boldsymbol{\alpha})$ be as in the previous section, and write $f(\boldsymbol{\alpha})=f(\boldsymbol{\alpha};P,R)$. Further, put $s=t+2u+v$, and let

$$R_s(\mathbf{n})=\int_{\mathbb{T}^{k+1}}F(\boldsymbol{\alpha})^tf(\boldsymbol{\alpha})^{2u+v}e(-\boldsymbol{\alpha}\cdot\mathbf{n})\,d\boldsymbol{\alpha}.$$

Then we have $W_s(\mathbf{n},P)\ge R_s(\mathbf{n})$, so it suffices to obtain a lower bound for $R_s(\mathbf{n})$. We dissect $\mathbb{T}^{k+1}$ into major and minor arcs as follows. Recalling the notation of Theorem 3, define

$$\mathfrak{m}=\mathfrak{m}_{1/2}\quad\text{and}\quad\mathfrak{M}=\mathbb{T}^{k+1}\setminus\mathfrak{m}.$$

We take

$$t=(k+1)^2,\quad u=\left[\frac{7}{3}k^2\log k+\frac{5}{3}k^2\log\log k+6k^2\right],\quad\text{and}\quad v=\left[\frac{\Delta_u}{\sigma_1(k)}\right]+1,$$

where Δ_u is as in Theorem 2 and $\sigma_1(k)$ is as in Corollary 3.1. A simple calculation shows that $v\ll k^2$, and hence

$$s=\frac{14}{3}k^2\log k+\frac{10}{3}k^2\log\log k+O(k^2).$$

On applying the aforementioned theorem and corollary, we find that

$$\begin{aligned}\int_{\mathfrak{m}}|F(\boldsymbol{\alpha})|^t|f(\boldsymbol{\alpha})|^{2u+v}d\boldsymbol{\alpha} &\ll P^{2t}\sup_{\boldsymbol{\alpha}\in\mathfrak{m}}|f(\boldsymbol{\alpha})|^v\int_{\mathbb{T}^{k+1}}|f(\boldsymbol{\alpha})|^{2u}d\boldsymbol{\alpha}\\ &\ll P^{2s-k(k+1)-\delta}\end{aligned}\tag{9.1}$$

for some $\delta>0$, since $\Delta_u<v\sigma_1(k)$. Thus it remains to deal with the major arcs.

When $(q,a_0,\ldots,a_k)=1$, define

$$\mathfrak{M}(q,\mathbf{a})=\{\boldsymbol{\alpha}\in\mathbb{T}^{k+1}:|q\alpha_i-a_i|\le P^{1/2-k}R^k\ (0\le i\le k)\},\tag{9.2}$$

so that

$$\mathfrak{M}=\bigcup_{\substack{1\le a_0,\ldots,a_k\le q\le P^{1/2}R^{k+1}\\(q,a_0,\ldots,a_k)=1}}\mathfrak{M}(q,\mathbf{a}).$$

It is a simple exercise to show that the $\mathfrak{M}(q, \mathbf{a})$ are pairwise disjoint. On recalling the notation of the previous section, we can record the following major arc approximation for $F(\boldsymbol{\alpha})$.

Lemma 9.1. *Suppose that $\boldsymbol{\alpha} \in \mathfrak{M}(q, \mathbf{a})$, and write $\beta_i = \alpha_i - a_i/q$. Then one has*

$$F(\boldsymbol{\alpha}) - q^{-2} S(q, \mathbf{a}) v(\boldsymbol{\beta}) \ll P^{3/2+\varepsilon}.$$

Proof. This follows immediately from Lemma 8.1, together with (9.2). □

The following estimates for $S(q, \mathbf{a})$, $v(\boldsymbol{\beta})$, and $w(\boldsymbol{\beta})$ are essentially immediate from the work of Arkhipov, Karatsuba, and Chubarikov [2].

Lemma 9.2. *Whenever $(q, a_0, \dots, a_k) = 1$, we have*

$$S(q, \mathbf{a}) \ll q^{2-1/k+\varepsilon}.$$

Proof. This follows easily from [2], Lemma II.8, on recalling standard divisor function estimates. □

Lemma 9.3. *One has*

$$v(\boldsymbol{\beta}) \ll P^2 (1 + P^k(|\beta_0| + \cdots + |\beta_k|))^{-1/k}$$

and

$$w(\boldsymbol{\beta}) \ll P^2 (1 + P^k(|\beta_0| + \cdots + |\beta_k|))^{-1/k}.$$

Proof. The first estimate follows from [2], Lemma II.2, on making a change of variable, and the second follows in a similar manner (see the comment in the proof of [13], Lemma 8.6) on noting that $\rho(\log \gamma / \log R) \asymp 1$ and is decreasing for $R \le \gamma \le P$. □

We now use the information contained in the above lemmata to prune back to a very thin set of major arcs on which $f(\boldsymbol{\alpha})$ can be suitably approximated. Specifically, let W be a parameter at our disposal, and recall the definition of $\mathfrak{N}(q, \mathbf{a})$ given in (8.2). Further, let

$$\mathfrak{N} = \bigcup_{\substack{1 \le a_0, \dots, a_k \le q \le W \\ (q, a_0, \dots, a_k) = 1}} \mathfrak{N}(q, \mathbf{a}). \tag{9.3}$$

We have the following result, which is closely analogous to [13], Lemma 9.2.

Lemma 9.4. *If t is an integer with $t \ge (k+1)^2$, then one has*

$$\int_{\mathfrak{M}} |F(\boldsymbol{\alpha})|^t \, d\boldsymbol{\alpha} \ll P^{2t - k(k+1)}$$

and

$$\int_{\mathfrak{M} \setminus \mathfrak{N}} |F(\boldsymbol{\alpha})|^t \, d\boldsymbol{\alpha} \ll W^{-\sigma} P^{2t - k(k+1)}$$

for some $\sigma > 0$.

Proof. When $\boldsymbol{\alpha} \in \mathfrak{M}(q, \mathbf{a})$, we have by Lemma 9.1 that

$$F(\boldsymbol{\alpha})^t - V(\boldsymbol{\alpha})^t \ll \left(P^{(3/2+\varepsilon)}\right)^t + P^{3/2+\varepsilon} |V(\boldsymbol{\alpha})|^{t-1}, \tag{9.4}$$

and the proof now follows the argument of Wooley [13], Lemma 9.2, employing our Lemma 9.2 together with the estimate

$$v(\boldsymbol{\beta}) \ll P^2 \prod_{i=0}^{k} (1 + P^k|\beta_i|)^{-1/k(k+1)},$$

which is immediate from Lemma 9.3. □

On making a trivial estimate for $f(\boldsymbol{\alpha})$, it follows directly from Lemma 9.4 that

$$\int_{\mathfrak{M}\setminus\mathfrak{N}} |F(\boldsymbol{\alpha})|^t |f(\boldsymbol{\alpha})|^{2u+v}\, d\boldsymbol{\alpha} \ll W^{-\sigma} P^{2s-k(k+1)} \tag{9.5}$$

for some $\sigma > 0$, so it suffices to deal with the pruned major arcs $\mathfrak{N}$. When $\boldsymbol{\alpha} \in \mathfrak{N}(q, \mathbf{a})$, we have by Lemma 8.4 that

$$f(\boldsymbol{\alpha})^{2u+v} - W(\boldsymbol{\alpha})^{2u+v} \ll \left(\frac{q^2P^2W^2}{\log P}\right)^{2u+v} + \frac{q^2P^2W^2}{\log P} |W(\boldsymbol{\alpha})|^{2u+v-1},$$

where

$$W(\boldsymbol{\alpha}) = W(\boldsymbol{\alpha}; q, \mathbf{a}) = q^{-2} S(q, \mathbf{a}) w(\boldsymbol{\beta}) \quad \text{and} \quad \beta_i = \alpha_i - a_i/q.$$

On combining this with (9.4) and recalling the definition of $\mathfrak{N}$, we find that

$$\int_{\mathfrak{N}} F(\boldsymbol{\alpha})^t f(\boldsymbol{\alpha})^{2u+v}\, d\boldsymbol{\alpha} = \int_{\mathfrak{N}} V(\boldsymbol{\alpha})^t W(\boldsymbol{\alpha})^{2u+v}\, d\boldsymbol{\alpha} + O(P^{2s-k(k+1)} (\log P)^{-\delta})$$

for some $\delta > 0$, provided that W is chosen to be a suitably small power of $\log P$.

Now let

$$S(q) = \sum_{\substack{1 \le a_0, \ldots, a_k \le q \\ (q, a_0, \ldots, a_k) = 1}} (q^{-2} S(q, \mathbf{a}))^s e\left(\frac{-a_0 n_0 - \cdots - a_k n_k}{q}\right),$$

$$\mathfrak{S}(\mathbf{n}, P) = \sum_{q \le W} S(q),$$

and

$$\mathfrak{S}(\mathbf{n}) = \sum_{q=1}^{\infty} S(q).$$

Notice that by Lemma 9.2 we have $S(q) \ll q^{k+1-s/k+\varepsilon}$, whence

$$\mathfrak{S}(\mathbf{n}) \ll 1 \quad \text{and} \quad \mathfrak{S}(\mathbf{n}) - \mathfrak{S}(\mathbf{n}, P) \ll P^{-\delta}$$

for some $\delta > 0$, provided that $s \ge (k+1)^2$. Further, let

$$J(\mathbf{n}, P) = \int_{\mathcal{B}(P)} v(\boldsymbol{\beta})^t w(\boldsymbol{\beta})^{2u+v} e(-\boldsymbol{\beta} \cdot \mathbf{n})\, d\boldsymbol{\beta},$$

where

$$\mathcal{B}(P) = [-WP^{-k}, WP^{-k}]^{k+1},$$

and put

$$J(\mathbf{n}) = \int_{\mathbb{R}^{k+1}} v(\boldsymbol{\beta})^t w(\boldsymbol{\beta})^{2u+v} e(-\boldsymbol{\beta}\cdot\mathbf{n})\, d\boldsymbol{\beta}.$$

Then when $s \geq (k+1)^2$, it follows easily from Lemmata 9.2 and 9.3 that

$$J(\mathbf{n}) \ll P^{2s-k(k+1)}$$

and

$$\sum_{1\leq q\leq P^{1/2+\varepsilon}} |S(q)||J(\mathbf{n}) - J(\mathbf{n}, P)| \ll P^{2s-k(k+1)}(\log P)^{-\delta}$$

for some $\delta > 0$. Combining these observations, we find that

$$\int_{\mathfrak{N}} F(\boldsymbol{\alpha})^t f(\boldsymbol{\alpha})^{2u+v} e(-\boldsymbol{\alpha}\cdot\mathbf{n})\, d\boldsymbol{\alpha} = \mathfrak{S}(\mathbf{n})J(\mathbf{n}) + O(P^{2s-k(k+1)}(\log P)^{-\delta}) \tag{9.6}$$

for some $\delta > 0$, again provided that W is a sufficiently small power of $\log P$. The singular integral $J(\mathbf{n})$ and the singular series $\mathfrak{S}(\mathbf{n})$ require further analysis.

Lemma 9.5. *Suppose that $s \geq (k+1)^2$, and fix real numbers $\mu_0, \ldots, \mu_k$ with the property that the system* (1.7) *has a non-singular real solution with $0 < \eta_i, \xi_i < 1$. Then there exists a positive number $\delta' = \delta'(s, k, \boldsymbol{\mu})$ such that, whenever*

$$|n_j - P^k\mu_j| < \delta' P^k \quad (0 \leq j \leq k)$$

and P is sufficiently large, one has

$$J(\mathbf{n}) \gg P^{2s-k(k+1)}.$$

Proof. After a change of variables, we have

$$J(\mathbf{n}) = P^{2s-k(k+1)} \int_{\mathbb{R}^{k+1}} \int_{\mathcal{B}} T(\boldsymbol{\gamma}, \boldsymbol{\nu})\, e\left(\sum_{j=0}^{k} \beta_j(\phi_j(\boldsymbol{\gamma}, \boldsymbol{\nu}) - \mu_j + \delta_j)\right) d\boldsymbol{\gamma}\, d\boldsymbol{\nu}\, d\boldsymbol{\beta},$$

where

$$\mathcal{B} = [0,1]^{2t} \times [R/P, 1]^{4u+2v},$$

$$T(\boldsymbol{\gamma}, \boldsymbol{\nu}) = \prod_{i=t+1}^{s} \rho\left(\frac{\log P\gamma_i}{\log R}\right) \rho\left(\frac{\log P\nu_i}{\log R}\right), \tag{9.7}$$

$$\phi_j(\boldsymbol{\gamma}, \boldsymbol{\nu}) = \gamma_1^{k-j}\nu_1^j + \cdots + \gamma_s^{k-j}\nu_s^j,$$

and where $|\delta_j| \leq \delta'$ for each j. Notice that $(\boldsymbol{\eta}, \boldsymbol{\xi})$ is contained in $\mathcal{B}$ for P sufficiently large. Now let

$$\mathcal{S}(t_0, \ldots, t_k) = \{(\boldsymbol{\gamma}, \boldsymbol{\nu}) \in \mathcal{B} : \phi_j(\boldsymbol{\gamma}, \boldsymbol{\nu}) - \mu_j + \delta_j = t_j \ (0 \leq j \leq k)\},$$

so that

$$J(\mathbf{n}) = P^{2s-k(k+1)} \int_{\mathbb{R}^{k+1}} \int_{\mathcal{C}} \int_{\mathcal{S}(t_0,\ldots,t_k)} T(\boldsymbol{\gamma}, \boldsymbol{\nu})\, e(\beta_0 t_0 + \cdots + \beta_k t_k)\, d\mathcal{S}(\mathbf{t})\, d\mathbf{t}\, d\boldsymbol{\beta},$$

where $\mathcal{C} \subset \mathbb{R}^{k+1}$. Since $(\boldsymbol{\eta}, \boldsymbol{\xi}) \in \mathcal{B}$, we see that $\mathcal{C}$ contains a neighborhood of $(\delta_0, \ldots, \delta_k)$ and hence contains the origin when δ' is sufficiently small. Thus after $k+1$ applications of Fourier's Integral Theorem (see for example Davenport [5]) we obtain

$$J(\mathbf{n}) = P^{2s-k(k+1)} \int_{\mathcal{S}(\mathbf{0})} T(\boldsymbol{\gamma}, \boldsymbol{\nu})\, d\mathcal{S}(\mathbf{0}).$$

Now, for δ' sufficiently small, the implicit function theorem shows that $\mathcal{S}(\mathbf{0})$ is a space of dimension $2s-k-1$ with positive $(2s-k-1)$-dimensional measure, and the lemma follows on noting that $T(\boldsymbol{\gamma}, \boldsymbol{\nu}) \gg 1$ for $R/P \le \gamma, \nu \le 1$. $\square$

It remains to deal with p-adic solubility considerations and hence to obtain a lower bound for the singular series $\mathfrak{S}(\mathbf{n})$.

Lemma 9.6. *The function $S(q)$ is multiplicative.*

Proof. By [2], Lemma II.4, one has $S(qr, \mathbf{a}) = S(q, r^{k-1}\mathbf{a})S(r, q^{k-1}\mathbf{a})$ whenever $(q, r) = 1$, and the result now follows by a standard argument. $\square$

For each prime p, write

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

Whenever $s \ge (k+1)^2$ one finds using Lemmata 9.2 and 9.6 that

$$\mathfrak{S}(\mathbf{n}) = \prod_p \sigma(p) \tag{9.8}$$

and that there exists a constant $C(k)$ such that

$$\frac{1}{2} \le \prod_{p > C(k)} \sigma(p) \le \frac{3}{2}. \tag{9.9}$$

Hence it remains to deal with small primes. Let $M_{\mathbf{n}}(q)$ denote the number of solutions of the system of congruences

$$x_1^{k-j} y_1^j + \cdots + x_s^{k-j} y_s^j \equiv n_j \pmod{q} \quad (0 \le j \le k).$$

Lemma 9.7. *One has*

$$\sum_{d|q} S(d) = q^{k+1-2s} M_{\mathbf{n}}(q).$$

Proof. By the orthogonality of the additive characters modulo q, one has

$$M_{\mathbf{n}}(q) = \frac{1}{q^{k+1}} \sum_{r_0=1}^{q} \cdots \sum_{r_k=1}^{q} \left(S(q, \mathbf{r})\right)^s e\left(-(\mathbf{r} \cdot \mathbf{n})/q\right).$$

Now on writing $d = q/(q, r_0, \ldots, r_k)$ and $a_i = r_i d/q$ we obtain

$$M_{\mathbf{n}}(q) = \frac{1}{q^{k+1}} \sum_{d|q} \sum_{\substack{1 \le a_0, \ldots, a_k \le d \\ (d, a_0, \ldots, a_k) = 1}} (q/d)^{2s} \left(S(d, \mathbf{a})\right)^s e\left(-(\mathbf{a} \cdot \mathbf{n})/d\right),$$

and the result follows. $\square$

We therefore have

$$\sigma(p)=\lim_{h\to\infty}p^{h(k+1-2s)}M_{\mathbf{n}}(p^h), \tag{9.10}$$

so to show that $\mathfrak{S}(\mathbf{n})\gg 1$ it suffices to obtain a suitable lower bound for $M_{\mathbf{n}}(p^h)$. In order to deduce this from the existence of non-singular p-adic solutions to (1.6), we need a version of Hensel's Lemma. In what follows, we write $|\cdot|_p$ for the usual p-adic valuation, normalized so that $|p|_p=p^{-1}$.

Lemma 9.8. *Let $\psi_1,\dots,\psi_r$ be polynomials in $\mathbb{Z}_p[x_1,\dots,x_r]$ with Jacobian $\Delta(\boldsymbol{\psi};\mathbf{x})$, and suppose that $\mathbf{a}\in\mathbb{Z}_p^r$ satisfies*

$$|\psi_j(\mathbf{a})|_p<|\Delta(\boldsymbol{\psi};\mathbf{a})|_p^2\quad(1\le j\le r).$$

Then there exists a unique $\mathbf{b}\in\mathbb{Z}_p^r$ such that

$$\psi_j(\mathbf{b})=0\quad(1\le j\le r)\quad\text{and}\quad|b_i-a_i|_p<p^{-1}|\Delta(\boldsymbol{\psi};\mathbf{a})|_p\quad(1\le i\le r).$$

Proof. This is Proposition 5.20 of Greenberg [6] with $R=\mathbb{Z}_p$. □

Lemma 9.9. *Suppose that the system (1.6) has a non-singular p-adic solution. Then there exists an integer $u=u(p)$ such that whenever $h\ge u$ one has*

$$M_{\mathbf{n}}(p^h)\ge p^{(h-u)(2s-k-1)}.$$

Proof. We relabel the variables by writing

$$(z_1,\dots,z_{2s})=(x_1,\dots,x_s,y_1,\dots,y_s),$$

and let $\mathbf{a}=(a_1,\dots,a_{2s})$ be a non-singular p-adic solution of (1.6). Then there exist indices $i_0,\dots,i_k$ such that $\Delta(\boldsymbol{\psi};a_{i_0},\dots,a_{i_k})\ne 0$, so we can find an integer u such that

$$|\Delta(\boldsymbol{\psi};a_{i_0},\dots,a_{i_k})|_p^2=p^{1-u}>0.$$

Now suppose that $h\ge u$. For $i\notin\{i_0,\dots,i_k\}$, choose integers w_i with $w_i\equiv a_i \pmod{p^u}$, and write $v_i=a_i$ for $i=i_0,\dots,i_k$ and $v_i=w_i$ otherwise. Then on writing

$$\psi_j(\mathbf{z})=\psi_j(\mathbf{x},\mathbf{y})=x_1^{k-j}y_1^j+\dots+x_s^{k-j}y_s^j-n_j$$

for $0\le j\le k$, we see that

$$\psi_j(\mathbf{v})\equiv\psi_j(\mathbf{a})\equiv 0\pmod{p^u},$$

and hence

$$|\psi_j(\mathbf{v})|_p\le p^{-u}<|\Delta(\boldsymbol{\psi};v_{i_0},\dots,v_{i_k})|_p^2.$$

Now if $h\ge u$ then there are $p^{(h-u)(2s-k-1)}$ possible choices for the w_i modulo p^h. Moreover, for any fixed choice we may regard ψ_j as a polynomial in the $k+1$ variables $z_{i_0},\dots,z_{i_k}$ after substituting $z_i=w_i$ on the remaining indices. Thus for each admissible choice of $\mathbf{w}$ we may apply Lemma 9.8 to obtain integers $b_{i_0},\dots,b_{i_k}$ such that $\psi_j(\mathbf{b};\mathbf{w})\equiv 0\pmod{p^h}$ for each j, whence the lemma follows. □

Now by (9.10) and Lemma 9.9 we have $\sigma(p)\ge p^{u(k+1-2s)}$ for all primes p, so on combining this with (9.8) and (9.9) we see that $\mathfrak{S}(\mathbf{n})\gg 1$. Hence the proof of Theorem 4 is complete upon recalling Lemma 9.5, together with (9.1), (9.5), and (9.6).

10. Lines on Additive Equations

We now establish Theorems 5 and 6 by proceeding much as in the previous section. Before embarking on the circle method, however, we need to make some preliminary observations.

Lemma 10.1. *Suppose that* $(\mathbf{x},\mathbf{y}) \in \mathbb{R}^{2s}$ *is a solution of* (1.12)*, and let* $a, b, c,$ *and* d *be arbitrary real numbers. Then* $(a\mathbf{x}+b\mathbf{y}, c\mathbf{x}+d\mathbf{y})$ *is also a solution.*

Proof. For $0 \le j \le k$, write

$$A_j(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{s} c_i(ax_i+by_i)^{k-j}(cx_i+dy_i)^j.$$

Then by the binomial theorem we have for each j that

$$\begin{aligned} A_j(\mathbf{x},\mathbf{y}) &= \sum_{i=1}^{s} c_i \sum_{r=0}^{k-j} \binom{k-j}{r}(ax_i)^{k-j-r}(by_i)^r \sum_{s=0}^{j}\binom{j}{s}(cx_i)^{j-s}(dy_i)^s \\ &= \sum_{r=0}^{k-j}\sum_{s=0}^{j}\binom{k-j}{r}\binom{j}{s} a^{k-j-r}b^r c^{j-s}d^s \sum_{i=1}^{s} c_i x_i^{k-(r+s)} y_i^{r+s}, \end{aligned}$$

and the lemma follows. □

Lemma 10.2. *Suppose that the system of equations* (1.12) *has a non-singular real solution* $(\boldsymbol{\eta},\boldsymbol{\xi})$*. Then we can find a non-singular real solution* $(\boldsymbol{\eta}',\boldsymbol{\xi}')$ *such that* η_i' *and* ξ_i' *are nonzero for each* i*.*

Proof. For $0 \le j \le k$, let

$$\psi_j(\mathbf{x},\mathbf{y}) = c_1x_1^{k-j}y_1^j + \cdots + c_sx_s^{k-j}y_s^j,$$

and write $(z_0,\ldots,z_{2s-1}) = (x_1,\ldots,x_s,y_1,\ldots,y_s)$. Then by rearranging variables, we may write the given real solution as $(\boldsymbol{\eta},\boldsymbol{\xi}) = (\gamma_0,\ldots,\gamma_{2s-1})$, where

$$\det\left(\frac{\partial\psi_j}{\partial z_i}(\boldsymbol{\gamma})\right)_{0\le i,j\le k} \ne 0.$$

Hence by using the Implicit Function Theorem as in the proof of [13], Lemma 6.2, we see that there exists a $(2s-k-1)$-dimensional neighborhood T_0 of $(\gamma_{k+1},\ldots,\gamma_{2s-1})$ and a function $\phi : T_0 \to \mathbb{R}^{k+1}$ such that $\boldsymbol{\gamma} = (\phi(\mathbf{w}),\mathbf{w})$ is a solution of (1.12) whenever $\mathbf{w} \in T_0$. Thus by choosing $\mathbf{w}$ with $|w_i-\gamma_i|$ sufficiently small for $k+1 \le i \le 2s-1$, we may assume that $\boldsymbol{\gamma}$ is a non-singular solution whose last $2s-k-1$ coordinates are nonzero. Moreover, a simple calculation shows that at most two of the remaining η_i and at most two of the remaining ξ_i are zero and that either η_i or ξ_i is nonzero for every i. In particular, when $s \ge 5$, there is some i for which $\eta_i\xi_i \ne 0$. Now let

$$b = \min\{|\eta_i/\xi_i| : \eta_i\xi_i \ne 0\} \quad \text{and} \quad c = \min\{|\xi_i/\eta_i| : \eta_i\xi_i \ne 0\},$$

and take $b' < \frac{1}{2}b$ and $c' < \frac{1}{2}c$. Then by Lemma 10.1 we see that $(\boldsymbol{\eta}',\boldsymbol{\xi}')$ is a solution of (1.12), where $\boldsymbol{\eta}' = \boldsymbol{\eta} + b'\boldsymbol{\xi}$ and $\boldsymbol{\xi}' = c'\boldsymbol{\eta} + \boldsymbol{\xi}$, and it is easy to check that η_i' and ξ_i' are nonzero for each i. The non-singularity follows by continuity on choosing b' and c' sufficiently small. □

By Lemma 10.2 we may henceforth suppose that the system (1.12) has a non-singular real solution $(\boldsymbol{\eta}, \boldsymbol{\xi})$ with η_i and ξ_i nonzero for all i, and by homogeneity we can re-scale to ensure that $0 < |\eta_i|, |\xi_i| < \frac{1}{2}$. For each i, write

$$\eta_i^+ = \eta_i + \tfrac{1}{2}|\eta_i| \quad \text{and} \quad \eta_i^- = \eta_i - \tfrac{1}{2}|\eta_i|$$

and

$$\xi_i^+ = \xi_i + \tfrac{1}{2}|\xi_i| \quad \text{and} \quad \xi_i^- = \xi_i - \tfrac{1}{2}|\xi_i|.$$

Now let P be a large positive number, put $R = P^\eta$ with $\eta \leq \eta_0(\varepsilon, k)$, and let $c_1, \ldots, c_s$ be nonzero integers. Throughout this section, the implicit constants arising in our analysis may depend on $c_1, \ldots, c_s$ and on the real solution $(\boldsymbol{\eta}, \boldsymbol{\xi})$. We define the exponential sums

$$F_i(\boldsymbol{\alpha}) = \sum_{\eta_i^- P < x \leq \eta_i^+ P} \sum_{\xi_i^- P < y \leq \xi_i^+ P} e(c_i(\alpha_0 x^k + \alpha_1 x^{k-1} y + \cdots + \alpha_k y^k))$$

and

$$f_i(\boldsymbol{\alpha}) = \sum_{\substack{\eta_i^- P < x \leq \eta_i^+ P \\ |x| \in \mathcal{A}(P,R)}} \sum_{\substack{\xi_i^- P < y \leq \xi_i^+ P \\ |y| \in \mathcal{A}(P,R)}} e(c_i(\alpha_0 x^k + \alpha_1 x^{k-1} y + \cdots + \alpha_k y^k)).$$

Further, write $s = t + 2u + v$ and define

$$\mathcal{F}(\boldsymbol{\alpha}) = \prod_{i=1}^{t} F_i(\boldsymbol{\alpha}) \quad \text{and} \quad \mathcal{G}(\boldsymbol{\alpha}) = \prod_{i=t+1}^{s} f_i(\boldsymbol{\alpha}).$$

Finally, let

$$R_s(P) = \int_{\mathbb{T}^{k+1}} \mathcal{F}(\boldsymbol{\alpha}) \mathcal{G}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha}.$$

Then we have $N_s(P) \geq R_s(P)$, so to prove Theorem 5 it suffices to obtain a lower bound for $R_s(P)$. We dissect $\mathbb{T}^{k+1}$ into major and minor arcs as follows. Write $c = \max |c_i|$ and $X = cP^{1/2} R^{k+1}$, and define

$$\mathfrak{M} = \bigcup_{\substack{1 \leq a_0, \ldots, a_k \leq q \leq X \\ (q, a_0, \ldots, a_k) = 1}} \mathfrak{M}(q, \mathbf{a}),$$

where

$$\mathfrak{M}(q, \mathbf{a}) = \{ \boldsymbol{\alpha} \in \mathbb{T}^{k+1} : |q\alpha_i - a_i| \leq P^{1/2-k} R^k \ (0 \leq i \leq k) \},$$

and put $\mathfrak{m} = \mathbb{T}^{k+1} \setminus \mathfrak{M}$. As before, it is easily seen that the $\mathfrak{M}(q, \mathbf{a})$ are disjoint

Lemma 10.3. *Whenever $\boldsymbol{\alpha} \in \mathfrak{m}$, one has $c_i \boldsymbol{\alpha} \in \mathbf{m}_{1/2}$. Moreover,*

$$\sup_{\boldsymbol{\alpha} \in \mathfrak{m}} |f_i(\boldsymbol{\alpha})| \ll P^{2 - \sigma_1(k) + \varepsilon},$$

where $\sigma_1(k)$ is as in Corollary 3.1.

Proof. Suppose that $\boldsymbol{\alpha} \in \mathfrak{m}$ and that $|c_i \alpha_j q - a_j| \leq P^{1/2-k} R^k$ for $0 \leq j \leq k$, where $q \in \mathbb{N}$, $a_j \in \mathbb{Z}$, and $(q, a_0, \ldots, a_k) = 1$. Then one has

$$\left| \alpha_j - \frac{a_j}{c_i q} \right| \leq \frac{P^{1/2-k} R^k}{|c_i| q} \quad (0 \leq j \leq k),$$

so on writing

$$d = (c_i, a_0, \dots, a_k), \quad a'_j = \frac{|c_i| a_j}{c_i d}, \quad \text{and} \quad q' = \frac{|c_i| q}{d},$$

we see that

$$\left| \alpha_j - \frac{a'_j}{q'} \right| \le \frac{P^{1/2-k} R^k}{q'd} \quad (0 \le j \le k),$$

so we must have $cq \ge q' > cP^{1/2}R^{k+1}$ and hence $q > P^{1/2}R^{k+1}$. Thus $c_i\boldsymbol{\alpha} \in \mathbf{m}_{1/2}$. The second assertion now follows on recalling the remark at the end of Section 7 and noting that we may replace α_j by $-\alpha_j$ as needed so that our sums are over positive integers. □

As in the previous section, we take

$$t = (k+1)^2, \quad u = \left[\frac{7}{3}k^2 \log k + \frac{5}{3}k^2 \log\log k + 6k^2 \right], \quad \text{and} \quad v = \left[\frac{\Delta_u}{\sigma_1(k)} \right] + 1,$$

where Δ_u is as in Theorem 2 and $\sigma_1(k)$ is as in Corollary 3.1. Then by Hölder's inequality and a change of variables we obtain

$$\begin{aligned} \int_{\mathfrak{m}} |\mathcal{F}(\boldsymbol{\alpha})\mathcal{G}(\boldsymbol{\alpha})| \, d\boldsymbol{\alpha} &\ll P^{2t+v(2-\sigma_1+\varepsilon)} \prod_{i=t+1}^{t+2u} \left(\int_{\mathbb{T}^{k+1}} |f_i(\boldsymbol{\alpha})|^{2u} d\boldsymbol{\alpha} \right)^{1/2u} \\ &\ll P^{2s-k(k+1)-\delta} \end{aligned} \tag{10.1}$$

for some $\delta > 0$, since $\Delta_u < v\sigma_1(k)$. Thus it remains to deal with the major arcs.

Recalling the notation of the previous section, we define $S_i(q, \mathbf{a}) = S(q, c_i\mathbf{a})$,

$$v_i(\boldsymbol{\beta}) = \int_{\eta_i^- P}^{\eta_i^+ P} \int_{\xi_i^- P}^{\xi_i^+ P} e(c_i(\beta_0\gamma^k + \beta_1\gamma^{k-1}\nu + \cdots + \beta_k\nu^k)) \, d\gamma \, d\nu,$$

and $V_i(\boldsymbol{\alpha}) = q^{-2}S_i(q, \mathbf{a})v_i(\boldsymbol{\alpha} - \mathbf{a}/q)$ for $\boldsymbol{\alpha} \in \mathfrak{M}(q, \mathbf{a})$. Further, we define the pruned major arcs $\mathfrak{N}$ as in the previous section using (8.2) and (9.3), again with W a suitable power of $\log P$. Finally, write

$$w_i(\boldsymbol{\beta}) = \int_{\eta_i^- P}^{\eta_i^+ P} \int_{\xi_i^- P}^{\xi_i^+ P} \rho\left(\frac{\log\gamma}{\log R}\right) \rho\left(\frac{\log\nu}{\log R}\right) e(c_i(\beta_0\gamma^k + \beta_1\gamma^{k-1}\nu + \cdots + \beta_k\nu^k)) \, d\gamma \, d\nu$$

and $W_i(\boldsymbol{\alpha}) = q^{-2}S_i(q, \mathbf{a})w_i(\boldsymbol{\alpha} - \mathbf{a}/q)$ for $\boldsymbol{\alpha} \in \mathfrak{N}(q, \mathbf{a})$. The next several lemmas are simple adaptations of the corresponding results in the previous section.

Lemma 10.4. *When $\boldsymbol{\alpha} \in \mathfrak{M}(q, \mathbf{a})$, one has*

$$F_i(\boldsymbol{\alpha}) - V_i(\boldsymbol{\alpha}) \ll P^{3/2+\varepsilon},$$

and when $\boldsymbol{\alpha} \in \mathfrak{N}(q, \mathbf{a})$, one has

$$f_i(\boldsymbol{\alpha}) - W_i(\boldsymbol{\alpha}) \ll \frac{q^2P^2W^2}{\log P}.$$

Proof. These estimates follow by making trivial modifications in the arguments of Lemmata 8.1 and 8.4, respectively. □

Lemma 10.5. *Whenever* $(q, a_0, \dots, a_k) = 1$, *we have*

$$S_i(q, \mathbf{a}) \ll q^{2-1/k+\varepsilon}.$$

Proof. Put $d_i = (q, c_i)$. Then by Lemma 9.2 we have

$$S_i(q, \mathbf{a}) = d_i^2 S(q/d_i, c_i\mathbf{a}/d_i) \ll d_i^{1/k} q^{2-1/k+\varepsilon} \ll_c q^{2-1/k+\varepsilon},$$

as required. □

Lemma 10.6. *One has*

$$v_i(\boldsymbol{\beta}) \ll P^2(1 + P^k(|\beta_0| + \cdots + |\beta_k|))^{-1/k}$$

and

$$w_i(\boldsymbol{\beta}) \ll P^2(1 + P^k(|\beta_0| + \cdots + |\beta_k|))^{-1/k}.$$

Proof. The argument is identical to the proof of Lemma 9.3. □

Lemma 10.7. *If t is an integer with $t \geq (k+1)^2$, then one has*

$$\int_{\mathfrak{M}} |F_i(\boldsymbol{\alpha})|^t \, d\boldsymbol{\alpha} \ll P^{2t-k(k+1)} \tag{10.2}$$

and

$$\int_{\mathfrak{M}\setminus\mathfrak{N}} |F_i(\boldsymbol{\alpha})|^t \, d\boldsymbol{\alpha} \ll W^{-\sigma} P^{2t-k(k+1)} \tag{10.3}$$

for some $\sigma > 0$.

Proof. The result follows as in Lemma 9.4 on using Lemmata 10.4, 10.5, and 10.6 in place of the corresponding results in the previous section. □

Once again, Lemma 10.7, together with (10.1), allows us to focus attention on the pruned major arcs $\mathfrak{N}$. Let

$$S(q) = \sum_{\substack{1 \leq a_0, \dots, a_k \leq q \\ (q, a_0, \dots, a_k) = 1}} q^{-2s} \prod_{i=1}^{s} S_i(q, \mathbf{a}),$$

$$\mathfrak{S}(P) = \sum_{q \leq X} S(q), \quad \text{and} \quad \mathfrak{S} = \sum_{q=1}^{\infty} S(q).$$

Again we have $S(q) \ll q^{k+1-s/k+\varepsilon}$, and hence $\mathfrak{S} \ll 1$ and $\mathfrak{S} - \mathfrak{S}(P) \ll P^{-\delta}$ for some $\delta > 0$, provided that $s \geq (k+1)^2$. Further, let

$$J(P) = \int_{\mathcal{B}(P)} \prod_{i=1}^{t} v_i(\boldsymbol{\beta}) \prod_{i=t+1}^{s} w_i(\boldsymbol{\beta}) \, d\boldsymbol{\beta},$$

where $\mathcal{B}(P) = [-WP^{-k}, WP^{-k}]^{k+1}$, and put

$$J = \int_{\mathbb{R}^{k+1}} \prod_{i=1}^{t} v_i(\boldsymbol{\beta}) \prod_{i=t+1}^{s} w_i(\boldsymbol{\beta}) \, d\boldsymbol{\beta}.$$

Then when $s \geq (k+1)^2$, we have by Lemmata 10.5 and 10.6 that $J \ll P^{2s-k(k+1)}$ and
$$\sum_{1 \leq q \leq cP^{1/2+\varepsilon}} |S(q)||J - J(P)| \ll P^{2s-k(k+1)}(\log P)^{-\delta}.$$
for some $\delta > 0$. Thus, by employing standard arguments based on Lemmata 10.4, 10.5, and 10.6, we obtain
$$\int_{\mathfrak{N}} \mathcal{F}(\boldsymbol{\alpha})\mathcal{G}(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} = \mathfrak{S}J + O(P^{2s-k(k+1)}(\log P)^{-\delta}) \tag{10.4}$$
for some $\delta > 0$.

Lemma 10.8. *Whenever $s \geq (k+1)^2$ and P is sufficiently large, one has*
$$J \gg P^{2s-k(k+1)}.$$

Proof. By a change of variables, we find that
$$J = P^{2s-k(k+1)} \int_{\mathbb{R}^{k+1}} \int_{\mathcal{B}} T(\boldsymbol{\gamma}, \boldsymbol{\nu})\, e\left(\sum_{j=0}^{k} \beta_j \phi_j(\boldsymbol{\gamma}, \boldsymbol{\nu})\right) d\boldsymbol{\gamma}\, d\boldsymbol{\nu}\, d\boldsymbol{\beta},$$
where
$$\mathcal{B} = [\eta_1^-, \eta_1^+] \times \cdots \times [\eta_s^-, \eta_s^+] \times [\xi_1^-, \xi_1^+] \times \cdots \times [\xi_s^-, \xi_s^+],$$
$$\phi_j(\boldsymbol{\gamma}, \boldsymbol{\nu}) = c_1 \gamma_1^{k-j} \nu_1^j + \cdots + c_s \gamma_s^{k-j} \nu_s^j,$$
and where $T(\boldsymbol{\gamma}, \boldsymbol{\nu})$ is as in (9.7). Now let
$$\mathcal{S}(t_0, \ldots, t_k) = \{(\boldsymbol{\gamma}, \boldsymbol{\nu}) \in \mathcal{B} : \phi_j(\boldsymbol{\gamma}, \boldsymbol{\nu}) = t_j \ (0 \leq j \leq k)\},$$
so that
$$J = P^{2s-k(k+1)} \int_{\mathbb{R}^{k+1}} \int_{\mathcal{C}} \int_{\mathcal{S}(t_0,\ldots,t_k)} T(\boldsymbol{\gamma}, \boldsymbol{\nu})\, e(\beta_0 t_0 + \cdots + \beta_k t_k)\, d\mathcal{S}(\mathbf{t})\, d\mathbf{t}\, d\boldsymbol{\beta},$$
where $\mathcal{C} \subset \mathbb{R}^{k+1}$. Since $(\boldsymbol{\eta}, \boldsymbol{\xi}) \in \mathcal{B}$, we see that $\mathcal{C}$ contains a neighborhood of the origin, whence after $k+1$ applications of Fourier's Integral Theorem we obtain
$$J = P^{2s-k(k+1)} \int_{\mathcal{S}(\mathbf{0})} T(\boldsymbol{\gamma}, \boldsymbol{\nu})\, d\mathcal{S}(\mathbf{0}),$$
and the result follows as in the proof of Lemma 9.5. □

Lemma 10.9. *The function $S(q)$ is multiplicative.*

Proof. This is identical to the proof of Lemma 9.6. □

Whenever $s \geq (k+1)^2$ one finds using Lemma 10.9 that
$$\mathfrak{S} = \prod_p \sigma(p), \quad \text{where} \quad \sigma(p) = \sum_{h=0}^{\infty} S(p^h),$$
and that there exists a constant $C(k)$ such that
$$\frac{1}{2} \leq \prod_{p > C(k)} \sigma(p) \leq \frac{3}{2}.$$

Let $M(q)$ denote the number of solutions of the system of congruences

$$c_1x_1^{k-j}y_1^j+\cdots+c_sx_s^{k-j}y_s^j\equiv 0 \pmod q \quad (0\le j\le k).$$

Lemma 10.10. *One has*

$$\sum_{d|q}S(d)=q^{k+1-2s}M(q).$$

Proof. This is identical to the proof of Lemma 9.7. □

It follows that

$$\sigma(p)=\lim_{h\to\infty}p^{h(k+1-2s)}M(p^h),$$

so again to show that $\mathfrak{S}\gg 1$ it suffices to show that $M(p^h)\ge p^{(h-u)(2s-k-1)}$ for $p\le C(k)$, and this follows exactly as in the argument of Lemma 9.9. Hence the proof of Theorem 5 is complete on assembling (10.1), (10.3), and (10.4) and recalling Lemma 10.8.

In order to deduce Theorem 6, we need some additional observations.

Lemma 10.11. *Let $(\mathbf{x},\mathbf{y}),(\mathbf{x}',\mathbf{y}')\in\mathbb{Z}^{2s}$ be such that $(x_1,\dots,x_s)=1$. Then $\mathbf{x}t+\mathbf{y}$ and $\mathbf{x}'t+\mathbf{y}'$ parameterize the same line if and only if*

$$\mathbf{x}'=q\mathbf{x}\quad\textit{and}\quad\mathbf{y}'=\mathbf{y}+r\mathbf{x}$$

for some integers q and r with $q\neq 0$.

Proof. First suppose that $\mathbf{x}'=q\mathbf{x}$ and $\mathbf{y}'=\mathbf{y}+r\mathbf{x}$ for some integers q and r with $q\neq 0$. Then one has

$$\mathbf{x}t+\mathbf{y}=\mathbf{x}'\left(\frac{t-r}{q}\right)+\mathbf{y}'\quad\text{and}\quad\mathbf{x}'t+\mathbf{y}'=\mathbf{x}(qt+r)+\mathbf{y},$$

so the two lines are identical. Conversely, suppose that the two lines are the same. By taking $t=0$ on the line $\mathbf{x}'t+\mathbf{y}'$, we see that there exists t_1 such that $\mathbf{y}'=\mathbf{x}t_1+\mathbf{y}$, and then by taking $t=1$ we find that there exists t_2 such that $\mathbf{x}'+\mathbf{y}'=\mathbf{x}t_2+\mathbf{y}$ and hence $\mathbf{x}'=(t_2-t_1)\mathbf{x}$. Moreover, the condition $(x_1,\dots,x_s)=1$ implies that t_1 and t_2 are distinct integers, and this completes the proof. □

Now let $R_s(P,d)$ denote the number of solutions of (1.12) counted by $R_s(P)$ for which $(x_1,\dots,x_s)=d$. Further, let $N_s'(P)$ denote the number of solutions counted by $N_s(P)$ for which $(x_1,\dots,x_s)=1$ and $1\le y_1\le |x_1|$. The following estimate will be useful when d is large.

Lemma 10.12. *One has*

$$R_s(P,d)\ll\frac{P^{2s-k(k+1)}}{d^2}.$$

Proof. Consider a solution $(\mathbf{x},\mathbf{y})$ counted by $R_s(P,d)$. Since x_{s-1} and x_s each have d as a divisor, the number of possible choices for x_{s-1},y_{s-1},x_s, and y_s is at most $P^2(P/d)^2$. Given such a choice, the number of possibilities for the remaining variables is

$$\int_{\mathbb{T}^{k+1}}\left(\prod_{i=1}^{t}F_i(\boldsymbol{\alpha})\prod_{i=t+1}^{s-2}f_i(\boldsymbol{\alpha})\right)e(\boldsymbol{\alpha}\cdot\mathbf{m})\,d\boldsymbol{\alpha},$$

where $m_j = c_{s-1}x_{s-1}^{k-j}y_{s-1}^j + c_s x_s^{k-j}y_s^j$, and thus

$$R_s(P,d) \ll \frac{P^4}{d^2}\int_{\mathbb{T}^{k+1}} \prod_{i=1}^{t} |F_i(\boldsymbol{\alpha})| \prod_{i=t+1}^{s-2} |f_i(\boldsymbol{\alpha})|\, d\boldsymbol{\alpha}.$$

The lemma now follows by dissecting $\mathbb{T}^{k+1}$ into major and minor arcs and using (10.1) and (10.2). □

We can now complete the proof of Theorem 6.

Proof of Theorem 6. Define an equivalence relation on the set of solutions to (1.12) by writing $(\mathbf{x},\mathbf{y}) \sim (\mathbf{x}',\mathbf{y}')$ whenever $\mathbf{x}t+\mathbf{y}$ and $\mathbf{x}'t+\mathbf{y}'$ define the same line. Thus we need a lower bound for the number of equivalence classes.

Let $N_s(P,Q,d)$ be the number of solutions of (1.12) with

$$\mathbf{x} \in \mathcal{B}P, \quad \mathbf{y} \in \mathcal{C}Q, \quad \text{and} \quad (x_1,\dots,x_s) = d,$$

where

$$\mathcal{B} = [\eta_1^-,\eta_1^+] \times \cdots \times [\eta_s^-,\eta_s^+] \quad \text{and} \quad \mathcal{C} = [\xi_1^-,\xi_1^+] \times \cdots \times [\xi_s^-,\xi_s^+].$$

Then the solutions counted by $R_s(P,d)$ and $N_s(P/d,P,1)$ are in bijective correspondence, and Lemma 10.11 shows that two solutions $(\mathbf{x},\mathbf{y})$ and $(\mathbf{x}',\mathbf{y}')$ counted by $N_s(P/d,P,1)$ are equivalent if and only if $\mathbf{x}=\mathbf{x}'$ and $\mathbf{y}-\mathbf{y}'=r\mathbf{x}$ for some integer r. Then since

$$|x_1| \geq \frac{P}{\tau d} \quad \text{and} \quad |y_1 - y_1'| \leq P,$$

where $\tau = 2/|\eta_1|$, we see that each equivalence class contains at most τd members counted by $N_s(P/d,P,1)$. Moreover, Lemma 10.1 allows us to map each equivalence class to a solution counted by $N_s'(P)$, and Lemma 10.11 shows that this map is injective. Thus we see that

$$R_s(P,d) = N_s(P/d,P,1) \leq \tau d\, N_s'(P). \tag{10.5}$$

Now let D be a parameter at our disposal. Since any two solutions counted by $N_s'(P)$ represent distinct equivalence classes, we have by (10.5) that

$$\sum_{d \leq D} R_s(P,d) \leq \tau D^2 N_s'(P) = \tau D^2 L_s(P).$$

Thus by Lemma 10.12 there exist positive constants γ_1 and γ_2 such that

$$\gamma_1 P^{2s-k(k+1)} \leq R_s(P) \leq \tau D^2 L_s(P) + \sum_{d>D}\left(\frac{\gamma_2 P^{2s-k(k+1)}}{d^2}\right)$$

for P sufficiently large, and hence we have

$$L_s(P) \geq \frac{P^{2s-k(k+1)}}{\tau D^2}\left(\gamma_1 - \frac{\gamma_2}{D}\right).$$

The theorem now follows on taking $D = 2\gamma_2/\gamma_1$.

References

[1] G. I. Arkhipov and A. A. Karatsuba, *A multidimensional analogue of Waring's problem*, Soviet Math. Dokl. **36** (1988), 75–77.

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

[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] H. Davenport, *Analytic methods for diophantine equations and diophantine inequalities*, Ann Arbor Publishers, 1962.

[6] M. J. Greenberg, *Lectures on forms in many variables*, New York-Amsterdam, 1969.

[7] A. Iarrobino, *The inverse system of a symbolic power, II: The Waring problem for forms*, J. Algebra **174** (1995), 1091–1110.

[8] D. J. Lewis and R. Schulze-Pillot, *Linear spaces on the intersection of cubic hypersurfaces*, Monatsh. Math. **97** (1984), 277–285.

[9] R. Miranda, *Linear systems of plane curves*, Notices Amer. Math. Soc. **46** (1999), no. 2, 192–202.

[10] S. T. Parsell, *The density of rational lines on cubic hypersurfaces*, Trans. Amer. Math. Soc. **352** (2000), 5045–5062.

[11] R. C. Vaughan, *A new iterative method in Waring's problem*, Acta Math. **162** (1989), 1–71.

[12] ———, *The Hardy-Littlewood method, second edition*, Cambridge University Press, Cambridge, 1997.

[13] T. D. Wooley, *On simultaneous additive equations II*, J. reine angew. Math. **419** (1991), 141–198.

[14] ———, *Large improvements in Waring's problem*, Ann. Math. **135** (1992), 131–164.

[15] ———, *Forms in many variables*, Analytic Number Theory: Proceedings of the 39th Taniguchi International Symposium, Kyoto, May, 1996 (Y. Motohashi, ed.), London Math. Soc. Lecture Note Ser. 247, Cambridge University Press, 1997, pp. 361–376.

[16] ———, *Linear spaces on cubic hypersurfaces and pairs of homogeneous cubic equations*, Bull. London Math. Soc. **29** (1997), 556–562.

[17] ———, *On exponential sums over smooth numbers*, J. reine angew. Math. **488** (1997), 79–140.

Department of Mathematics, Texas A&M University, College Station, Texas 77843-3368
E-mail address: `parsell@alum.mit.edu`