PAIRS OF ADDITIVE EQUATIONS OF SMALL DEGREE

SCOTT T. PARSELL

1. INTRODUCTION

The investigation of systems of additive diophantine equations has long been a favorite application of the Hardy-Littlewood method, and such investigations have played a large role in motivating various refinements of the method (see for example [5], [9], [10], [12]–[15], and [27]). For years, however, virtually all the attention was focused on systems in which the equations are all of the same degree. This pattern was finally broken in the early 1990s with the work of Wooley [35]–[37] on systems consisting of one cubic and one quadratic equation. Owing to this breakthrough and its subsequent refinements (see Wooley [40], [41]), the technology now exists to allow further study of simultaneous equations of differing degree. It is non-trivial, however, to obtain results that are close to optimal in the sense that they exploit the vast modern machinery of the circle method to its fullest extent, and this may partially explain why systems of higher degree (and systems consisting of more than two equations) have not yet been considered in detail.

The purpose of this paper is to obtain a collection of explicit results concerning pairs of equations of differing degree and, in doing so, to provide an array of estimates and a computational model for handling further cases. To be more specific, let k and n be integers with $k \geq n \geq 1$, and let $c_1, \ldots, c_s$ and $d_1, \ldots, d_s$ be integers. We consider the problem of determining conditions under which the system of equations

$$(1.1) \qquad \begin{aligned} c_1 x_1^k + \cdots + c_s x_s^k &= 0 \\ d_1 x_1^n + \cdots + d_s x_s^n &= 0 \end{aligned}$$

possesses a non-trivial integral solution. One obvious requirement is that the system have a non-trivial real solution and a non-trivial p-adic solution for every prime p. In fact, the success of the Hardy-Littlewood method depends explicitly upon good information concerning the density of such

solutions, and hence one must typically demonstrate the existence of non-singular solutions in each local field. Unfortunately, the local solubility problem for (1.1) tends to be quite hard. When $p > k^4 n^2$, Wooley [35] has shown that there exist non-trivial p-adic solutions to (1.1), provided only that $s > 2(k + n)$, and this latter bound is further shown to be best possible for infinitely many primes p whenever $n > 1$. The number of variables required for smaller primes in the recent work of Knapp [20], however, is often significantly larger than what would be required to handle the minor arcs in an application of the circle method. Moreover, it is generally difficult to guarantee that the solutions produced by the p-normalization methods of [20] and [35] are non-singular. When $k - n = 1$, a non-trivial singular solution satisfies linear relations that enable one to construct a non-trivial rational solution (see equation (6.9) of Wooley [37]), but this procedure would appear to have no obvious analogue when $k - n > 1$. We therefore focus our attention on determining how large s must be in order to establish a local-global principle for (1.1).

Write $\Gamma^*(k, n)$ for the least integer r such that, whenever $s \geq r$, the system (1.1) has a non-trivial p-adic solution for each prime p. In this notation, it follows from our earlier remarks that $\Gamma^*(k, n) \geq 2(k + n) + 1$ whenever $n > 1$. When considering global solubility, we mostly confine our attention to systems in which each of the variables $x_1, \ldots, x_s$ appears explicitly in both equations. When $k > n$, we call the system (1.1) *regular* if for each i one has $c_i d_i \neq 0$. When $k = n$, it is instead convenient to call the system regular if one has $c_i d_j \neq c_j d_i$ whenever $i \neq j$. Note that, for a given s, almost all systems (1.1) are regular, in the sense of natural density in $\mathbb{Z}^{2s}$. Define $G^*(k, n)$ to be the least integer r such that, whenever $s \geq r$, every regular system (1.1) having a non-singular real solution and a non-singular p-adic solution for all primes p has a non-trivial integral solution.

As mentioned above, the situation in which $k = n$ has attracted interest for quite some time, beginning with work of Davenport and Lewis [14] on pairs of cubics in 18 variables. Cook [12] showed that $G^*(2, 2) \leq 9$, while the bound of Davenport and Lewis for $G^*(3, 3)$ has been steadily reduced over the years by work of Cook [13], Vaughan [27], Baker and Brüdern [5], and finally Brüdern [9], who obtained $G^*(3, 3) \leq 14$ in a massive display of technical prowess. An obstruction to local solubility is already apparent here, as Davenport and Lewis [14] show that there are pairs of cubics in 15 variables possessing no non-trivial 7-adic solution. In attempting to bound $G^*(k, k)$, one is essentially able to consider Weyl sums over kth powers, for which good mean value estimates are available from the extensive work of Vaughan and Wooley [32], [34] on Waring's problem. As a result, one typically expects bounds for $G^*(k, k)$ that are about twice the size of the corresponding bound for $G(k)$ in Waring's problem, except in cases where

$G(k)$ is hugely inflated by a local obstruction (for example, when $k = 4$). While it is often challenging to establish the bound with exactly twice as many variables as used in Waring's problem (see [9] and [24]), the inclusion of just one additional variable usually reduces the problem to a straightforward exercise.

In order to handle pairs of equations of differing degree, one requires a distinctly two-dimensional approach involving exponential sums over polynomials of the shape $\alpha x^k + \beta x^n$, for which sharp mean value estimates are less readily available. This situation was first tackled by Wooley [37], who developed a version of Vaughan's iterative method [29] suitable for generating such estimates. This work initially produced the bound $G^*(3, 2) \leq 14$, which was later improved to $G^*(3, 2) \leq 13$ (see [41]) using the refinements of [40]. In fact, Wooley [36] was able to show additionally that $\Gamma^*(3, 2) = 11$, so in this case the existence of integral solutions to (1.1) when $s \geq 13$ is subject only to a real solubility hypothesis.

It is a straightforward exercise to generate bounds for other pairs of exponents by using the general method of Wooley [40]. In particular, one should be able to demonstrate with little difficulty that $G^*(k, n) \leq (2+o(1))k \log k$, for large $k \geq n$. This would compare essentially as expected with the results of [10] and [38], although with more work one may be able to obtain asymptotics that take into account the size of n relative to k. As in the current treatment of Waring's problem (see Vaughan and Wooley [32], [34]), there are various refinements that may be attempted in order to obtain good results for smaller exponents. As the amount of available technology associated with the Hardy-Littlewood method is nowadays quite substantial, one has many options for carrying out such refinements, and it is a non-trivial task to determine the optimal strategy in each case. The primary goal of this paper is to develop and implement a collection of strategies leading to bounds for $G^*(k, n)$ that are essentially the best attainable within the current scope of the circle method. Moreover, the exponential sum estimates obtained here may be applied to related problems. For example, one can deal with pairs of diophantine inequalities, in which one seeks to demonstrate that two forms with real coefficients take arbitrarily small values simultaneously at integral points. This problem has already been investigated by the author [22], [23] in the case of a cubic and quadratic form, and we intend to return to this application in a later paper. For now, we are content to record our bounds for $G^*(k, n)$, together with those previously known, in the following theorem.

Theorem 1.1. *One has $G^*(k, n) \leq F(k, n)$, where $F(k, n)$ is given by the entry in row k and column n of the following table.*

	1	2	3	4	5	6	7
3	10	13	14				
4	17	20	24	24			
5	30	31	32	36	34		
6	49	50	49	47	50	49	
7	66	72	70	65	64	66	67

We remark that the value of $G^*(2,1)$ is a matter of convention, since the Hasse principle for quadratic forms applies immediately when the second equation in (1.1) is substituted into the first. As previously indicated, the bound for $G^*(3,2)$ is due to Wooley [41], and the bound for $G^*(3,3)$ is due to Brüdern [9]. The conclusion that $G^*(5,5) \leq 34$ was recently established by Parsell and Wooley [24]. Finally, the bounds quoted for $G^*(k,k)$ in the cases $k = 4,6,7$ can be obtained by relatively routine arguments based on the estimates of Vaughan [29] and Vaughan and Wooley [32], [34]; we leave the details to the reader. It is possible that an application of the p-adic iteration method of [24] may yield a one-variable reduction in the latter three bounds but there are non-trivial issues to be addressed in carrying this out. Since we are primarily concerned here with technology applying to equations of differing degree, we do not pursue such refinements.

It is worth mentioning that the arguments used to establish Theorem 1.1 yield quantitative information in most cases, namely that the number of solutions of (1.1) lying in the box $[-P, P]^s$ is of order P^{s-k-n}. In the cases $(k,n) = (3,3)$ and $(5,5)$, the lower bound is reduced somewhat by the use of prime variables, whereas the ad-hoc methods applied in the cases $(k,n) = (3,1)$ and $(7,1)$ yield at best substantially weaker estimates. We note, however, that the expected density results can always be recovered by increasing the number of variables; in the latter two cases, it would suffice to use 11 and 75 variables, respectively.

The regularity condition on the system (1.1) can be relaxed somewhat (see for example [10], [24] for the case $k = n$), although a general result allowing for zero coefficients (or repeated coefficient ratios when $k = n$) can be rather awkward to state. The theorem below, analogous to those of Wooley [37], [41], gives an indication of the type of result that may be obtained for forms of differing degree.

Theorem 1.2. *Suppose that $k = 5$ and $n = 3$. Then the simultaneous equations (1.1) have infinitely many solutions in rational integers provided that*

   (a) *the cubic equation has at least 7 variables explicit,*
   (b) *the quintic equation has at least 17 variables explicit,*

(c) *the simultaneous equations* (1.1) *have a non-singular real solution and a non-singular p-adic solution for every prime p, and*

(d) *either*

    (i) $s \geq 32$, *or*

    (ii) *at least* 16 *of the* d_i *are zero, or*

    (iii) *at least* 6 *of the* c_i *are zero.*

While Theorem 1.2 applies only to the case $k = 5$ and $n = 3$, there is in principle no difficulty in proving similar results for other pairs of degrees, although the cases where both forms are of odd degree are certainly the simplest to handle. The particular numbers of zero coefficients allowed or required in the various conditions depend upon the number of variables required to ensure the solubility of single additive equations of degrees k and n. Further, these conditions must be balanced against the total number of variables required in (d)(i), since zero coefficients introduce common factors on the major arcs that may hinder the convergence of the singular series for smaller values of s. On the other hand, the conditions in Theorem 1.2 are somewhat redundant; for example, if condition (d)(i) holds, then conditions (a) and (b) are superfluous, and hence the regularity condition can be removed completely in this case. Finally, the hypotheses of the theorem can be altered to produce quantitative results if desired. In the situation of Theorem 1.2, for example, one obtains the expected density of solutions by replacing the numbers 7 and 17 in conditions (a) and (b) by $s - 15$ and $s - 5$, respectively, and deleting options (ii) and (iii) of condition (d). We leave it to the reader to formulate similar theorems in other cases of interest.

We begin by establishing our fundamental efficient differencing procedure in §2. Some special cases involving linear forms are disposed of in §3, after which we develop our collection of end-game procedures in §§4-5. The optimal strategies and resulting mean value estimates are then recorded in §6, and the proof of Theorem 1.1 is then accessible by an essentially routine application of the circle method, which we describe in §7. Finally, in §8, we sketch the modifications required to establish results such as Theorem 1.2, in which one allows for the presence of zero coefficients.

Throughout, the letters ε and η denote sufficiently small positive numbers. We take P to be a large real number depending at most on s, k, n, ε, η, and the coefficients c_i and d_i. The implicit constants in Vinogradov's and Landau's notation depend at most on this same list of parameters, unless otherwise indicated. As usual, we write $e(z)$ for $e^{2\pi i z}$, and we make frequent use of vector notation, writing for example $\mathbf{x} = (x_1, \ldots, x_s)$. We also write $\mathbb{T}^2$ for the unit square $[0,1]^2$.

In order to simplify our exposition, we adopt the following convention concerning the parameters ε and R. Whenever ε or R appears in a statement, either implicitly or explicitly, then we assert that for each $\varepsilon > 0$, there exists a positive number $\eta_0 = \eta_0(\varepsilon, s, k, n)$ such that the statement holds whenever $R = P^\eta$ with $0 < \eta \leq \eta_0$. The values of ε and η_0 may therefore change from statement to statement, and hence also the dependence of implicit constants on ε and η. Since our iterative methods will involve only a finite number of statements, there is no danger of losing control of the implicit constants. Finally, we use the symbol $\approx$ to indicate that constants and powers of R and P^ε are to be ignored.

The author is extremely grateful to Professor Wooley for many helpful discussions concerning the methods employed in this paper and, in particular, for supplying an unpublished manuscript containing estimates of the type considered in Lemma 5.2.

2. REPEATED EFFICIENT DIFFERENCING

In this section, we describe the efficient differencing process that will be used to generate our mean value estimates. The essential analysis here is due to Wooley [40], so our main task is to write down the estimates of that paper, specialized to the case of two equations, in a form suitable for describing our iterative schemes. For convenience, we make use of symmetric difference polynomials, and this requires a slight modification in the analysis of [40], along the lines of Vaughan and Wooley [32], [34].

Let $j \leq k$ be a positive integer, and let $K_1, \ldots, K_j$ be integers satisfying $n \leq K_i \leq k$ and $1/K_1 + \cdots + 1/K_j \leq 1$. We write $\mathbf{K}_j$ for the vector $(K_1, \ldots, K_j)$ and define the efficient differencing operator $\Delta_{j,\mathbf{K}}^*$ recursively as follows. For any integer K, let

$$\Delta_{1,K}^*(f(x); h; m) = (f(x + hm^K) - f(x)),$$

and then define

$$\Delta_{j+1,\mathbf{K}_{j+1}}^*(f(x); h_1, \ldots, h_{j+1}; m_1, \ldots, m_{j+1})$$
$$= \Delta_{1,K_{j+1}}^*(\Delta_{j,\mathbf{K}_j}^*(f(x); h_1, \ldots, h_j; m_1, \ldots, m_j); h_{j+1}; m_{j+1}),$$

with the convention that $\Delta_{0,K}^*(f(x); h; m) = f(x)$. Now, when $0 \leq j \leq k$, write

$$\psi_{i,j}(z; h_1, \ldots, h_j; m_1, \ldots, m_j) = \Delta_{j,\mathbf{K}_j}^*(f(z), 2h_1, \ldots, 2h_j; m_1, \ldots, m_j),$$

where $f(z) = (z - h_1 m_1^{K_1} - \cdots - h_j m_j^{K_j})^i$. Let $\phi_j = \phi_j(s, k, n, \mathbf{K})$ be real numbers satisfying $0 \leq \phi_j \leq 1/K_j$, and write

$$P_j = 2^j P, \quad M_j = P^{\phi_j}, \quad H_j = P_j M_j^{-K_j}, \quad Q_j = P_j(M_1 \cdots M_j)^{-1}.$$

We adopt the convention that $Q_0 = P_0 = P$ and further write

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

For given non-zero integers $m_1, \ldots, m_j$, we also introduce the notation

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

In our applications, it will often be the case that $K_1 = \cdots = K_j = k$. In this situation, the polynomials

$$(2.1) \qquad \Psi_{i,j}(z; \mathbf{h}; \mathbf{m}) = \frac{\psi_{i,j}(z; \mathbf{h}; \mathbf{m})}{D_{i,j}(\mathbf{m})}$$

have integer coefficients for $i = k, n$ and are sometimes more convenient to work with than the $\psi_{i,j}$. Write $\boldsymbol{\alpha} = (\alpha_k, \alpha_n)$, and define the exponential sum

$$(2.2) \qquad F_j(\boldsymbol{\alpha}) = \sum_{m_1, \ldots, m_j} \sum_{h_1, \ldots, h_{j-1}} \left| \sum_{h_j} G_j(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m}) \right|^2,$$

where

$$(2.3) \qquad G_j(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m}) = \sum_{z} e(\alpha_k \Psi_{k,j}(z; \mathbf{h}; \mathbf{m}) + \alpha_n \Psi_{n,j}(z; \mathbf{h}; \mathbf{m})),$$

and where the summations range over

$$(2.4) \qquad M_i < m_i \le M_i R, \quad 1 \le h_i \le 2^{j-1-i} H_i, \quad \text{and} \quad 1 \le z \le P_j.$$

Further, let

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

denote the set of R-smooth numbers of size at most P, define

$$(2.5) \qquad f(\boldsymbol{\alpha}) = f(\boldsymbol{\alpha}; P, R) = \sum_{x \in \mathcal{A}(P,R)} e(\alpha_k x^k + \alpha_n x^n),$$

and write $f_j(\boldsymbol{\alpha}) = f(\boldsymbol{\alpha}; Q_j, R)$. We say that λ_s is a permissible exponent if one has the estimate

$$(2.6) \qquad \int_{\mathbb{T}^2} |f(\boldsymbol{\alpha})|^{2s} \, d\boldsymbol{\alpha} \ll P^{\lambda_s + \varepsilon},$$

where the roles of ε and R are as indicated at the end of §1.

We need to consider some auxiliary exponential sums in order to handle the more general situation in which the coefficients of $\Psi_{i,j}$ are not integral. First of all, write

$$\widetilde{f}_j(\boldsymbol{\alpha}; \mathbf{m}) = \sum_{x \in \mathcal{A}(Q_j, R)} e(\alpha_k D_{k,j}(\mathbf{m})x^k + \alpha_n D_{n,j}(\mathbf{m})x^n).$$

Next, write

$$\widetilde{G}_j(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m}) = \sum_z e(\alpha_k \psi_{k,j}(z; \mathbf{h}; \mathbf{m}) + \alpha_n \psi_{n,j}(z; \mathbf{h}; \mathbf{m})),$$

and define

$$(2.7) \qquad \widetilde{F}_j(\boldsymbol{\alpha}; \mathbf{m}) = \begin{cases} \displaystyle\sum_{h_1,\ldots,h_{j-1}} \left| \sum_{h_j} \widetilde{G}_j(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m}) \right|^2 & \text{if } 1 \leq j \leq n, \\ \displaystyle\sum_{h_1,\ldots,h_j} \widetilde{G}_j(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m}) & \text{if } n < j \leq k. \end{cases}$$

Finally, put

$$\mathcal{F}_j^{(s)}(\boldsymbol{\alpha}) = \sum_{m_1,\ldots,m_j} \widetilde{F}_j(\boldsymbol{\alpha}; \mathbf{m}) |\widetilde{f}_j(\boldsymbol{\alpha}; \mathbf{m})|^{2s}.$$

Here the variables $\mathbf{m}$, $\mathbf{h}$, and z in the summations range over the intervals (2.4).

For brevity, we write $r_j = 2$ when $0 \leq j < n$ and $r_j = 1$ when $n \leq j < k$, and we also set $\omega_j = K_j - n$ for $1 \leq j \leq n$ and $\omega_j = 0$ for $n < j \leq k$. The following two estimates, based on the method of Wooley [40], are fundamental to our efficient differencing process.

Lemma 2.1. *Suppose that λ_s is a permissible exponent. Then one has*

$$\int_{\mathbb{T}^2} F_0(\boldsymbol{\alpha})^2 |f_0(\boldsymbol{\alpha})|^{2s} \, d\boldsymbol{\alpha} \ll P^\varepsilon M_1^{2s+\omega_1-1} \left(P^2 M_1 Q_1^{\lambda_s} + \int_{\mathbb{T}^2} \mathcal{F}_1^{(s)}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \right).$$

Proof. This essentially follows from Lemma 5.1 of Wooley [40], except that here we are using symmetric difference polynomials. The argument of [40], Lemma 4.1, applies without any changes, so we are able to conclude that

$$\int_{\mathbb{T}^2} F_0(\boldsymbol{\alpha})^2 |f_0(\boldsymbol{\alpha})|^{2s} \, d\boldsymbol{\alpha} \ll P^\varepsilon M_1^{2s+\omega_1-1} T(P, Q_1, R),$$

where $T(P, Q_1, R)$ is the number of solutions of the system

$$(2.8) \qquad z_1^i + z_2^i - w_1^i - w_2^i = m^i \sum_{l=1}^s (u_l^i - v_l^i) \qquad (i = k, n),$$

with the variables satisfying

$$(2.9) \qquad u_l, v_l \in \mathcal{A}(\tfrac{1}{2} Q_1, R), \qquad M_1 < m \leq M_1 R,$$

$$1 \leq z_l, w_l \leq P \qquad \text{and} \qquad z_l \equiv w_l \pmod{m^{K_1}}.$$

We may now write

$$T(P, Q_1, R) = U_0 + U_1,$$

where U_0 denotes the number of solutions of (2.8) with $z_1 = w_1$ or $z_2 = w_2$ and U_1 denotes the number of solutions with $z_1 \neq w_1$ and $z_2 \neq w_2$. Then by applying the Cauchy-Schwarz inequality to the underlying integral of exponential sums as in the proof of [40], Lemma 4.2, we find that if $T(P, Q_1, R) \ll U_0$ then

$$U_0 \ll P^{2+\varepsilon} M_1 Q_1^{\lambda_s}.$$

For a non-diagonal solution counted by U_1, we can write

$$z_1 = w_1 + h_1 m^{K_1} \quad \text{and} \quad z_2 = w_2 + h_2 m^{K_1}$$

for some integers h_1 and h_2 satisfying

$$1 \leq |h_l| \leq \tfrac{1}{2} H_1 \qquad (l = 1, 2).$$

Now on writing $x_l = z_l + w_l$ for $l = 1, 2$, we find that

$$x_l + h_l m^{K_1} = 2 z_l \quad \text{and} \quad x_l - h_l m^{K_1} = 2 w_l,$$

so it follows that U_1 is bounded above by the number of solutions of the system

$$\eta_1 \psi_{i,1}(x_1; h_1; m) + \eta_2 \psi_{i,1}(x_2; h_2; m) = m^i \sum_{l=1}^{s} 2^i (u_l^i - v_l^i) \qquad (i = k, n)$$

with $\eta_1, \eta_2 \in \{\pm 1\}$, $1 \leq x_l \leq 2P$, and $1 \leq h_l \leq \tfrac{1}{2} H_1$, and with the remaining variables as in (2.9). On considering the underlying diophantine equations, it follows that

$$U_1 \ll \sum_{M_1 < m \leq M_1 R} \int_{\mathbb{T}^2} \widetilde{F}_1(\boldsymbol{\alpha}; m) |\widetilde{f}_1(\boldsymbol{\alpha}; m)|^{2s} \, d\boldsymbol{\alpha},$$

and this completes the proof of the lemma. $\qquad \square$

Lemma 2.2. *Suppose that $1 \leq j < k$ and $0 < t < 2s$ and that λ_t and λ_{2s-t} are permissible exponents. Then whenever $j \neq n$, one has*

$$\int_{\mathbb{T}^2} \mathcal{F}_j^{(s)}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \ll P^{\varepsilon} H_j^{r-1} \left(Q_j^{\lambda_t} \widetilde{H}_j \widetilde{M}_j M_{j+1}^{4s-2t+\omega_{j+1}-1} T_{j+1} \right)^{1/2},$$

where r denotes the common value of r_{j-1} and r_j, and where

$$(2.10) \qquad T_{j+1} = P^r \widetilde{H}_j \widetilde{M}_{j+1} Q_{j+1}^{\lambda_{2s-t}} + \int_{\mathbb{T}^2} \mathcal{F}_{j+1}^{(2s-t)}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha}.$$

Further, if λ_s is a permissible exponent, then one has

$$\int_{\mathbb{T}^2} \mathcal{F}_n^{(s)}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \ll P^{\varepsilon} H_n M_{n+1}^{2s-1} \left(P \widetilde{H}_n \widetilde{M}_{n+1} Q_{n+1}^{\lambda_s} + \int_{\mathbb{T}^2} \mathcal{F}_{n+1}^{(s)}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \right).$$

Proof. First suppose that $j \neq n$, write $r = r_{j-1} = r_j$, and suppose that $0 < t < 2s$. By two applications of the Cauchy-Schwarz inequalities, we find that

$$\int_{\mathbb{T}^2} \mathcal{F}_j^{(s)}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \leq \left(\widetilde{M}_j Q_j^{\lambda_t + \varepsilon} \sum_{m_1, \ldots, m_j} \int_{\mathbb{T}^2} |\widetilde{F}_j(\boldsymbol{\alpha}; \mathbf{m})^2 \widetilde{f}_j(\boldsymbol{\alpha}; \mathbf{m})^{4s - 2t}| \, d\boldsymbol{\alpha} \right)^{1/2},$$

and two more applications of Cauchy's inequality yield the estimate

$$|\widetilde{F}_j(\boldsymbol{\alpha}; \mathbf{m})|^2 \leq H_j^{2r-2} \widetilde{H}_j \sum_{h_1, \ldots, h_j} |\widetilde{G}_j(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m})|^{2r},$$

where the ranges of $\mathbf{m}$ and $\mathbf{h}$ are given by (2.4). We therefore conclude that

$$(2.11) \qquad \int_{\mathbb{T}^2} \mathcal{F}_j^{(s)}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \ll P^\varepsilon H_j^{r-1} \left(\widetilde{M}_j \widetilde{H}_j Q_j^{\lambda_t} S(P, Q_j, R) \right)^{1/2},$$

where $S(P, Q_j, R)$ is the number of solutions of the system

$$\sum_{l=1}^{r} \left(\psi_{i,j}(z_l; \mathbf{h}; \mathbf{m}) - \psi_{i,j}(w_l; \mathbf{h}; \mathbf{m}) \right) = D_{i,j}(\mathbf{m}) \sum_{l=1}^{2s-t} (u_l^i - v_l^i) \qquad (i = k, n),$$

with $\mathbf{h}$ and $\mathbf{m}$ satisfying (2.4), and with

$$(2.12) \qquad 1 \leq z_l, w_l \leq P_j \qquad \text{and} \qquad u_l, v_l \in \mathcal{A}(Q_j, R).$$

Now the argument of Wooley [40], Lemma 4.1, applies to give

$$(2.13) \qquad S(P, Q_j, R) \ll P^\varepsilon M_{j+1}^{4s - 2t + \omega_{j+1} - 1} T(P, Q_j, R),$$

where $T(P, Q_j, R)$ is the number of solutions of the system

$$(2.14) \qquad \begin{aligned} \sum_{l=1}^{r} &\left(\psi_{i,j}(z_l; \mathbf{h}; \mathbf{m}) - \psi_{i,j}(w_l; \mathbf{h}; \mathbf{m}) \right) \\ &= D_{i,j}(\mathbf{m}) q^i \sum_{l=1}^{2s-t} (u_l^i - v_l^i) \qquad (i = k, n) \end{aligned}$$

with

$$(2.15) \qquad u_l, v_l \in \mathcal{A}(\tfrac{1}{2} Q_{j+1}, R) \qquad \text{and} \qquad M_{j+1} < q \leq M_{j+1} R,$$

with $\mathbf{h}$ and $\mathbf{m}$ as in (2.4), with $\mathbf{z}$ and $\mathbf{w}$ as in (2.12), and additionally

$$z_l \equiv w_l \pmod{q^{K_{j+1}}} \qquad (1 \leq l \leq r).$$

At this point we follow the argument of Wooley [40], Lemma 4.2, as in the proof of Lemma 2.1 above. We have

$$T(P, Q_j, R) = U_0 + U_1,$$

where U_0 denotes the number of solutions of (2.14) with $z_l = w_l$ for some l and U_1 denotes the number of solutions with $z_l \neq w_l$ for each l. If U_0 dominates, we find as before that

$$(2.16) \qquad U_0 \ll P^{r+\varepsilon} \widetilde{H}_j \widetilde{M}_{j+1} Q_{j+1}^{\lambda_{2s-t}}.$$

Now for a solution counted by U_1, we can write

$$z_l = w_l + g_l q^{K_{j+1}} \qquad \text{with} \qquad 1 \le |g_l| \le P_j M_{j+1}^{-K_{j+1}} = \tfrac{1}{2} H_{j+1}.$$

Since the polynomial $\psi_{i,j}$ is homogeneous of degree i in the variables z, $h_1, \dots, h_j$, the change of variable $x_l = z_l + w_l$ shows that U_1 is bounded above by the number of solutions of the system

$$\sum_{l=1}^{r} \eta_l \psi_{i,j+1}(x_l; \mathbf{h}, g_l; \mathbf{m}, q) = D_{i,j}(\mathbf{m}) q^i \sum_{l=1}^{2s-t} 2^i (u_l^i - v_l^i) \qquad (i = k, n)$$

with $\mathbf{m}$ as in (2.4), with

$$\eta_l \in \{\pm 1\}, \qquad 1 \le x_l \le 2P_j, \qquad 1 \le h_i \le 2^{j-i} H_i, \qquad 1 \le g_l \le \tfrac{1}{2} H_{j+1},$$

and with the remaining variables as in (2.15). One therefore has

$$U_1 \ll \sum_{\mathbf{m}, q} \int_{\mathbb{T}^2} \widetilde{F}_{j+1}(\boldsymbol{\alpha}; \mathbf{m}, q) |\widetilde{f}_{j+1}(\boldsymbol{\alpha}; \mathbf{m}, q)|^{4s-2t} \, d\boldsymbol{\alpha},$$

where the ranges of summation are given by (2.4) and (2.15). The first part of the lemma now follows on recalling (2.11), (2.13), and (2.16) and considering the underlying diophantine equations.

Finally, we note that an application of Cauchy's inequality yields

$$\int_{\mathbb{T}^2} \mathcal{F}_n^{(s)}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \ll H_n S(P, Q_n, R),$$

where $S(P, Q_n, R)$ is as above (with $r = 1$ and $s = t$), and the second part of the lemma now follows by applying the argument ensuing from (2.13). $\square$

When $K_j = k$ for each j, we are able to divide through by $D_{i,j}(\mathbf{m})$ in the underlying equations and hence obtain Lemmas 2.1 and 2.2 with $\mathcal{F}_j^{(s)}(\boldsymbol{\alpha})$ replaced by $F_j(\boldsymbol{\alpha})|f_j(\boldsymbol{\alpha})|^{2s}$. Therefore, as in Vaughan and Wooley [32], [34], we write the inequalities embodied in these lemmas symbolically as

$$F_0^2 f_0^{2s} \longrightarrow F_1 f_1^{2s} \qquad \text{and} \qquad F_j f_j^{2s} \longrightarrow F_{j+1} f_{j+1}^{4s-2t}$$
$$\downarrow$$
$$f_j^{2t}$$

even though the sums F_j and f_j do not appear explicitly unless $K_1 = \cdots = K_j = k$.

Suppose we take a total of j differences using Lemmas 2.1 and 2.2. The values of t (and hence of s) on successive applications of Lemma 2.2 need not be the same, so we write $t_1, \ldots, t_{j-1}$ for the successive choices of t and put $s_{i+1} = 2s_i - t_i$ for $i = 1, \ldots, j - 1$, with the convention that $s_1 = s$. In our applications, we will take $K_1 = \cdots = K_{j-1} = k$ and allow only K_j to vary. This simplifies the computations considerably and represents no serious loss, since one typically expects that examining the underlying congruences with a reduced modulus would be useful primarily on the final difference, when we may be close to a diagonal situation.

We optimize the parameters $\phi_1, \ldots, \phi_j$ by first equating the two terms of the expression for T_j in (2.10). This yields a linear relation involving $\phi_1, \ldots, \phi_j$ that depends on how we estimate the integral in the second term; various possibilities for this "end-game" analysis will be discussed in §§4-5. Suppose now that $1 \le i < j$ and $i \ne n$, and that one has imposed constraints ensuring that

$$T_{i+1} \ll P^{r_i + \varepsilon} \widetilde{H}_i \widetilde{M}_{i+1} Q_{i+1}^{\lambda_{s_{i+1}}}.$$

Then by Lemma 2.2, one has

$$\int_{\mathbb{T}^2} \mathcal{F}_i^{(s_i)}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \ll P^{r_i/2 + \varepsilon} \widetilde{H}_i H_i^{r_i - 1} \widetilde{M}_i \left(M_{i+1}^{2s_{i+1} + \omega_{i+1}} Q_i^{\lambda_{t_i}} Q_{i+1}^{\lambda_{s_{i+1}}} \right)^{1/2},$$

so on equating the terms in the expression for T_i (or the terms in Lemma 2.1 when $i = 1$) we obtain the relation

$$P^{r_i} \widetilde{H}_{i-1} \widetilde{M}_i Q_i^{\lambda_{s_i}} \approx P^{r_i/2} \widetilde{H}_i H_i^{r_i - 1} \widetilde{M}_i \left(M_{i+1}^{2s_{i+1} + \omega_{i+1}} Q_i^{\lambda_{t_i}} Q_{i+1}^{\lambda_{s_{i+1}}} \right)^{1/2},$$

which simplifies to

$$P^{r_i} Q_i^{2\lambda_{s_i}} \approx H_i^{2r_i} M_{i+1}^{2s_{i+1} + \omega_{i+1}} Q_i^{\lambda_{t_i}} Q_{i+1}^{\lambda_{s_{i+1}}}.$$

When $i = n$, we take $s_n = t_n = s_{n+1}$, and one then finds from Lemma 2.2 that the above relation optimizes parameters in this case as well. On recalling that $K_i = k$ for $1 \le i < j$, we thus obtain the system of equations

(2.17) $$(\phi_1 + \cdots + \phi_{i-1})\mathcal{E}_i = r_i + \mathcal{E}_i - \phi_i(2r_i k + \mathcal{E}_i) + \phi_{i+1}\kappa_{i+1},$$

where

(2.18) $$\mathcal{E}_i = \lambda_{t_i} + \lambda_{s_{i+1}} - 2\lambda_{s_i} \qquad (1 \le i \le j - 1)$$

and

(2.19) $$\kappa_i = 2s_i + \omega_i - \lambda_{s_i} \qquad (2 \le i \le j).$$

The $j - 1$ equations (2.17), together with the equation arising from the estimation of the integral in (2.10) on the final difference, allow us to solve

successively for the optimal values of $\phi_1, \ldots, \phi_j$. Suppose that this final equation has the form

$$(2.20) \qquad \phi_j = \alpha_j(\gamma_j - \beta_j(\phi_1 + \cdots + \phi_{j-1})),$$

and define successively, for $i = j - 1, \ldots, 1$,

$$\gamma_i = r_i + \mathcal{E}_i + \kappa_{i+1}\alpha_{i+1}\gamma_{i+1},$$
$$\beta_i = \mathcal{E}_i + \kappa_{i+1}\alpha_{i+1}\beta_{i+1},$$
$$\alpha_i = (2r_i k + \beta_i)^{-1}.$$

Then a simple induction argument using (2.17) shows that

$$(2.21) \quad \phi_1 = \alpha_1\gamma_1 \quad \text{and} \quad \phi_i = \alpha_i(\gamma_i - \beta_i(\phi_1 + \cdots + \phi_{i-1})) \quad (2 \leq i \leq j).$$

The possible choices for α_j, β_j, and γ_j in (2.20) arising from various end-game strategies are discussed in §§4-5. Perhaps the simplest such strategy is to choose $\phi_j = 1/K_j$, which corresponds to taking

$$(2.22) \qquad \alpha_j = K_j^{-1}, \quad \beta_j = 0, \quad \text{and} \quad \gamma_j = 1$$

in (2.20). In this case, $H_j \ll 1$, so one may employ the trivial estimate

$$\widetilde{F}_j(\boldsymbol{\alpha}; \mathbf{m}) \ll \widetilde{H}_{j-1}(H_j P)^{r_j-1} \ll \widetilde{H}_{j-1}P^{r_j-1}$$

to deduce that

$$T_j \ll P^{r_j-1+\varepsilon}\widetilde{H}_{j-1}\widetilde{M}_j Q_j^{\lambda_{s_j}},$$

whence (2.21) holds with the values (2.22) inserted in (2.20). In certain situations, especially when n is small relative to k, this turns out to be the best we can do. In the next section, we describe some elementary alternatives to the circle method that are effective in some of these cases.

3. SYSTEMS WITH ONE FORM LINEAR

When differencing more than n times, there is only one non-trivial congruence implicit in (2.14), and thus one can no longer impose the non-singularity condition required to difference four variables effectively. As a result, we are able to difference only two variables at a time, just as in the one-dimensional analysis of [32] and [34]. It is therefore difficult to generate estimates of the quality that might be expected for $G^*(k, n)$ when n is small relative to k, since a large value of k would ordinarily call for more than n differences at certain stages of the iteration. As a consequence, we sometimes obtain better bounds for $G^*(k, n)$ when n is close to k than for $G^*(k, 1)$ or $G^*(k, 2)$. This is contrary to the expectation that systems of lower total degree should require fewer variables, but it seems to be a fundamental difficulty associated with the method.

Occasionally the mean value estimates generated by the methods of the previous section are so weak that certain ad-hoc approaches to solving the

system (1.1) yield results superior to those obtainable via the circle method. This phenomenon is especially evident in the case $n = 1$, and the purpose of this section is to describe how the bounds for $G^*(k, 1)$ in Theorem 1.1 are obtained in the cases $k = 3$ and $k = 7$. The arguments given in this section were suggested to the author by Professor Wooley, who has graciously allowed them to be recorded here.

First of all, we establish the bound $G^*(3, 1) \leq 10$. By solving the linear equation in (1.1) for x_{10} and substituting into the cubic, we obtain the equation

$$(3.1) \qquad d_{10}^3(c_1 x_1^3 + \cdots + c_9 x_9^3) - c_{10}(d_1 x_1 + \cdots + d_9 x_9)^3 = 0.$$

Let $C(\mathbf{x})$ denote the cubic form on the left-hand side of (3.1). In view of our local solubility hypotheses on (1.1), the form $C(\mathbf{x})$ has a non-trivial zero in $\mathbb{Q}_p$ for every prime p, so we can make use of Hooley's work on nonary cubic forms.

Lemma 3.1. *Let $f(\mathbf{x})$ be a non-singular cubic form in 9 variables with rational integral coefficients. Then the equation $f(\mathbf{x}) = 0$ has a non-trivial integral solution if $f(\mathbf{x})$ has a non-trivial zero in every p-adic field $\mathbb{Q}_p$.*

Proof. This is the main result of Hooley [18]. $\qquad\qquad\square$

In the case where $C(\mathbf{x})$ is non-singular, Lemma 3.1 ensures that the equation (3.1) has a non-trivial integral solution. Suppose now that $C(\mathbf{x})$ has a singular point $(\gamma_1, \ldots, \gamma_9) \neq \mathbf{0}$. By rearranging variables, we may assume that $\gamma_1 \neq 0$, and then by homogeneity we may further assume that $\gamma_1 \in \mathbb{Q}$. The singularity conditions are

$$d_{10}^3 c_i \gamma_i^2 = c_{10}(d_1 \gamma_1 + \cdots + d_9 \gamma_9)^2 d_i \quad (1 \leq i \leq 9),$$

and since $\gamma_1 \in \mathbb{Q}$, it follows that $(d_1 \gamma_1 + \cdots + d_9 \gamma_9)^2 \in \mathbb{Q}$. It is now immediate that each of $\gamma_2, \ldots, \gamma_9$ lies in a quadratic extension of $\mathbb{Q}$, and hence the equation $C(\mathbf{x}) = 0$ has a non-trivial solution in $\mathbb{Q}(\sqrt{r_1}, \ldots, \sqrt{r_8})$ for some $r_1, \ldots, r_8 \in \mathbb{Q}$. The following lemma, due essentially to Lewis [21], allows us to pull this back to a non-trivial solution in $\mathbb{Q}$.

Lemma 3.2. *Let K be a field, and suppose that $r_1, \ldots, r_t \in K$. If a cubic form $f(\mathbf{x})$ with coefficients in K has a non-trivial zero in $K(\sqrt{r_1}, \ldots, \sqrt{r_t})$, then it has a non-trivial zero in K.*

Proof. We first consider the case $t = 1$. Let $r \in K$, and suppose that $\mathbf{x} = \mathbf{a} + \mathbf{b}\sqrt{r}$ is a non-trivial zero of f in $K(\sqrt{r})$. Then we may assume that $\sqrt{r} \notin K$, that $\mathbf{b} \neq \mathbf{0}$, that $f(\mathbf{b}) \neq 0$, and that $\mathbf{a}$ and $\mathbf{b}$ are linearly independent over K, since otherwise f clearly has a non-trivial zero in K. By following the argument of Lewis [21], Lemma D (in which the result is proved for $r = -1$), we then find that $r\mathbf{a}f(\mathbf{b}) + \mathbf{b}f(\mathbf{a})$ is a non-trivial zero of f in K.

The general result is now obtained by applying the $t = 1$ case successively with K replaced by the various intermediate field extensions. $\qquad\square$

By combining Lemmas 3.1 and 3.2, we see that in any case (3.1) has a non-trivial rational solution and hence a non-trivial integral solution $(x_1, \ldots, x_9)$. Upon writing $y_i = d_{10}x_i$ for $1 \leq i \leq 9$ and $y_{10} = -(d_1 x_1 + \cdots + d_9 x_9)$, we then find that $(y_1, \ldots, y_{10})$ is a non-trivial integral solution of (1.1). We remark that a density result can be obtained for systems in 11 or more variables by the methods of §7. Indeed, the circle method just misses establishing the expected density with 10 variables, as the exponent $\lambda_3 = 6$ is permissible and a Weyl-type inequality saves nearly $P^{1/4}$ per variable.

We note here that the argument leading to (3.1) also shows that a quadratic-linear system of the type (1.1) has a non-trivial integral solution whenever $s \geq 6$, subject only to a real solubility hypothesis. This follows from Meyer's theorem that any indefinite quadratic form in five or more variables has a non-trivial integral zero. However, with the additional assumption of p-adic solubility, a non-trivial integral solution is guaranteed for any s by the Hasse-Minkowski theorem (see for example [7], §1.7.1). It is therefore a matter of convention to assign a value to $G^*(2, 1)$, but we note that $s = 3$ represents the smallest non-vacuous case.

Finally, in order to bound $G^*(7, 1)$, we make use of the following result on a single additive equation of degree 7.

Lemma 3.3. *If $s \geq 33$ and $c_1, \ldots, c_s$ are integers, then the equation*

$$c_1 x_1^7 + \cdots + c_s x_s^7 = 0$$

has a non-trivial integral solution.

Proof. The existence of non-trivial p-adic solutions for each prime p follows from the work of Dodson [16] whenever $s \geq 22$. The lemma then follows by a relatively straightforward adaptation of the methods of Vaughan and Wooley [31], [32]. $\qquad\square$

We now demonstrate how to solve the system (1.1) when $k = 7$, $n = 1$, and $s = 66$. By Lemma 3.3, one can find integers $x_1, \ldots, x_{33}$, not all zero, and $x_{34}, \ldots, x_{66}$, not all zero, such that

$$c_1 x_1^7 + \cdots + c_{33} x_{33}^7 = 0 \quad \text{and} \quad c_{34} x_{34}^7 + \cdots + c_{66} x_{66}^7 = 0.$$

Now write

$$M = d_1 x_1 + \cdots + d_{33} x_{33} \quad \text{and} \quad N = d_{34} x_{34} + \cdots + d_{66} x_{66}.$$

If $M = N = 0$, then $\mathbf{x} = (x_1, \ldots, x_{66})$ is a non-trivial solution to (1.1), and otherwise

$$\mathbf{y} = (N x_1, \ldots, N x_{33}, -M x_{34}, \ldots, -M x_{66})$$

is a non-trivial solution. We therefore conclude that $G^*(7,1) \leq 66$. In fact, we have demonstrated unconditional solubility when $s \geq 66$ as opposed to merely establishing a local-global principle, but the above argument fails to produce the expected density of solutions. One can recover the expected density, however, by using more variables and applying the circle method as in §7.

4. Estimates for Auxiliary Systems of Equations

We now begin to examine some more sophisticated methods for estimating the integral arising in the final application of Lemma 2.2. In this section, we consider estimates based on mean values of exponential sums closely related to $F_j(\boldsymbol{\alpha})$ and on direct analysis of the equations underlying the aforementioned integral. We first consider the exponential sum

$$\mathcal{G}_j^{(T)}(\boldsymbol{\alpha}) = \sum_{\mathbf{m}} |\widetilde{F}_j(\boldsymbol{\alpha}; \mathbf{m})|^T,$$

where $\widetilde{F}_j(\boldsymbol{\alpha}; \mathbf{m})$ is as in (2.7) and where the summation ranges over $m_1, \ldots, m_j$ satisfying (2.4). The following estimates will be particularly useful in our iterative process.

Lemma 4.1. *Suppose that $1 \leq j \leq k - 2$, and write $r = r_{j-1}$. Then one has*

$$\int_{\mathbb{T}^2} \mathcal{G}_j^{(2)}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \ll P^{r+\varepsilon} \widetilde{M}_j \widetilde{H}_j^r H_j^{r-1}.$$

Furthermore, if $1 \leq j \leq n - 1$, then one has

$$\int_{\mathbb{T}^2} \mathcal{G}_j^{(3)}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \ll P^{3+\varepsilon} \widetilde{M}_j \widetilde{H}_j^3 H_j^3.$$

Proof. First suppose that $j \leq n$, so that $r_{j-1} = 2$. By applying Cauchy's inequality and then Weyl differencing in (2.7), we find that for any fixed $\mathbf{m}$,

(4.1) $\widetilde{F}_j(\boldsymbol{\alpha}; \mathbf{m}) \ll H_j \left(P\widetilde{H}_j + |\Theta_j(\boldsymbol{\alpha}; \mathbf{m})| \right),$

where

$$\Theta_j(\boldsymbol{\alpha}; \mathbf{m}) = \sum_{\mathbf{h}} \sum_{u} \sum_{z \in \mathcal{I}(u)} e(\alpha_k \widetilde{\psi}_{k,j}(z; \mathbf{h}; \mathbf{m}; u) + \alpha_n \widetilde{\psi}_{n,j}(z; \mathbf{h}; \mathbf{m}; u)).$$

Here the summations range over $\mathbf{h}$ satisfying (2.4) and $1 \leq |u| < P_j$, and we have written

$$\widetilde{\psi}_{i,j}(z; \mathbf{h}; \mathbf{m}; u) = \psi_{i,j}(z + u; \mathbf{h}; \mathbf{m}) - \psi_{i,j}(z; \mathbf{h}; \mathbf{m})$$

and $\mathcal{I}(u) = [1 - u, P_j - u] \cap [1, P_j]$. Let us further write

$$\mathcal{J} = \int_{\mathbb{T}^2} \widetilde{F}_j(\boldsymbol{\alpha}; \mathbf{m})^2 \, d\boldsymbol{\alpha}.$$

We then have by (4.1) that

$$\mathcal{J} \ll H_j \left(P\widetilde{H}_j \int_{\mathbb{T}^2} \widetilde{F}_j(\boldsymbol{\alpha}; \mathbf{m}) \, d\boldsymbol{\alpha} + \int_{\mathbb{T}^2} |\Theta_j(\boldsymbol{\alpha}; \mathbf{m})| \widetilde{F}_j(\boldsymbol{\alpha}; \mathbf{m}) \, d\boldsymbol{\alpha} \right).$$

The integral in the first term on the right is bounded above by the number of solutions of the underlying equation of higher degree,

$$(4.2) \qquad \psi_{k,j}(z; h_1, \ldots, h_{j-1}, h_j; \mathbf{m}) = \psi_{k,j}(w; h_1, \ldots, h_{j-1}, g_j; \mathbf{m}),$$

with $1 \le z, w \le P_j$, with $h_1, \ldots, h_j$ as in (2.4), and with $1 \le g_j \le \frac{1}{2} H_j$. Moreover, the right hand side of (4.2) is non-zero and is divisible by g_j. Hence, after fixing $h_1, \ldots, h_j$ and z in $O(P\widetilde{H}_j)$ possible ways, we find that g_j is determined to $O(P^\varepsilon)$ by a divisor estimate, and w is then determined to $O(1)$ as the root of a non-trivial polynomial. We therefore have

$$\int_{\mathbb{T}^2} \widetilde{F}_j(\boldsymbol{\alpha}; \mathbf{m}) \, d\boldsymbol{\alpha} \ll P^{1+\varepsilon} \widetilde{H}_j,$$

so by applying the Cauchy-Schwarz inequality we deduce that

$$(4.3) \qquad \mathcal{J} \ll P^{2+\varepsilon} \widetilde{H}_j^2 H_j + H_j S_j(\mathbf{m})^{1/2} \mathcal{J}^{1/2},$$

where $S_j(\mathbf{m})$ is the number of solutions to the equation

$$(4.4) \qquad \widetilde{\psi}_{k,j}(z; \mathbf{h}; \mathbf{m}; u) = \widetilde{\psi}_{k,j}(w; \mathbf{g}; \mathbf{m}; v)$$

with z, w, and $\mathbf{h}$ as above, with $1 \le g_i \le 2^{j-1-i} H_j$, and with $1 \le u, v \le P_j$. We note that the right hand side of (4.4) is a non-zero multiple of $(g_1 \cdots g_j)v$. Since $j \le k - 2$, we find as in the above argument that a fixed choice of $\mathbf{h}$, u, and z determines $\mathbf{g}$, v, and w to $O(P^\varepsilon)$, whence

$$S_j(\mathbf{m}) \ll P^{2+\varepsilon} \widetilde{H}_j.$$

The first estimate of the lemma now follows in the case $j \le n$ by dividing into cases according to the dominant term on the right-hand side of (4.3) and then summing over $\mathbf{m}$. The case $j > n$ is easily established on recalling (2.7) and applying the argument surrounding (4.2).

The second estimate of the lemma follows by applying Hölder's inequality in (2.7) and invoking the main theorem of Parsell and Wooley [25], concerning nearly-diagonal behavior in symmetric systems of t equations in $2t + 2$ variables. $\qquad \square$

After differencing j times using Lemmas 2.1 and 2.2, we may apply Hölder's inequality to estimate the integral

$$\int_{\mathbb{T}^2} \mathcal{F}_j^{(s_j)}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha}$$

in terms of the moments of $\widetilde{F}_j$ and $\widetilde{f}_j$. Lemma 4.1 provides estimates for the second and third moments of $\widetilde{F}_j$, and estimates for higher moments can be established by similar arguments; however, the estimates obtainable for higher moments are typically too weak to be useful in light of the various alternative methods we will develop.

The estimate for the second moment of $\widetilde{F}_j$ is particularly useful. By applying the Cauchy-Schwarz inequalities, one finds that

$$\int_{\mathbb{T}^2} \mathcal{F}_j^{(s)}(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} \leq \left(\int_{\mathbb{T}^2} \mathcal{G}_j^{(2)}(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} \right)^{1/2} \left(\int_{\mathbb{T}^2} \sum_{\mathbf{m}} |\widetilde{f}_j(\boldsymbol{\alpha};\mathbf{m})|^{4s}\,d\boldsymbol{\alpha} \right)^{1/2},$$

where the summation is over $\mathbf{m}$ in the range given by (2.4). We denote an estimate of this form symbolically by

$$F_j f_j^{2s} \Longrightarrow (F_j^2)^{1/2}(f_j^{4s})^{1/2}.$$

Assuming that we have preliminary permissible exponents λ_{s_j} and λ_{2s_j} at our disposal (by applying the simple method described at the end of §2, for example), we deduce from Lemma 4.1 and a consideration of the underlying diophantine equations that

$$\int_{\mathbb{T}^2} \mathcal{F}_j^{(s_j)}(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} \ll P^\varepsilon \widetilde{M}_j (P^r \widetilde{H}_j^r H_j^{r-1} Q_j^{\lambda_{2s_j}})^{1/2},$$

where we have written $r = r_{j-1}$. Thus one optimizes parameters in (2.10) by taking

$$P^r \widetilde{H}_{j-1} \widetilde{M}_j Q_j^{\lambda_{s_j}} \approx \widetilde{M}_j (P^r \widetilde{H}_j^r H_j^{r-1} Q_j^{\lambda_{2s_j}})^{1/2},$$

which simplifies to

$$P^r \widetilde{H}_{j-1}^{2-r} Q_j^{2\lambda_{s_j}} \approx H_j^{2r-1} Q_j^{\lambda_{2s_j}}.$$

On writing $\delta_j = \lambda_{2s_j} - 2\lambda_{s_j}$, we find that

(4.5) $\phi_j((2r-1)K_j + \delta_j) = 1 + \delta_j + j(r-2) - (\delta_j + k(r-2))\Sigma_j,$

where $\Sigma_j = \phi_1 + \cdots + \phi_{j-1}$, so we take

$$\alpha_j = ((2r-1)K_j + \delta_j)^{-1}, \quad \beta_j = \delta_j + k(r-2), \quad \gamma_j = 1 + \delta_j + j(r-2)$$

in (2.20). Given parameters $s_1, \ldots, s_j, t_1, \ldots, t_{j-1}$, and K_j, we write A_j to denote the iterative process

$$F_0^2 f_0^{2s_1} \quad \to \quad F_1 f_1^{2s_1} \quad \to \quad F_2 f_2^{2s_2} \quad \to \cdots \to \quad F_j f_j^{2s_j} \quad \Rightarrow \quad (F_j^2)^{1/2}(f_j^{4s_j})^{1/2}$$
$$\downarrow \qquad\qquad\qquad \downarrow$$
$$f_1^{2t_1} \qquad\qquad\qquad f_{j-1}^{2t_{j-1}}$$

This process will typically be most useful for generating new permissible exponents towards the beginning of the iteration (*i.e.*, when s is relatively small).

Occasionally, it is useful (when $j \leq n - 1$) to employ both estimates of Lemma 4.1, in which case one considers a somewhat smaller (but still even integral) moment of $\widetilde{f}_j$. If one uses the $2u$th moment of $\widetilde{f}_j$, with $3s_j \leq 2u \leq 4s_j$, then Hölder's inequality can be applied to interpolate between the two estimates of Lemma 4.1 and hence to obtain an estimate for the $u/(u - s_j)$th moment of $\widetilde{F}_j$. By carrying out the above procedure to optimize parameters, one finds that

$$\alpha_j = (\delta_j - 3s_j K_j/u)^{-1}, \quad \beta_j = \delta_j, \quad \text{and} \quad \gamma_j = 1 + \delta_j - 3s_j/u,$$

where we have written $\delta_j = \lambda_{s_j} - (s_j/u)\lambda_u$. In this case, we denote the corresponding iterative process by A'_j.

Towards the middle of the iteration, it is often more effective to attempt a direct analysis of the pair of diophantine equations underlying the integral in (2.10). The remainder of this section is devoted to describing two methods for doing this. Unlike the process A_j, these methods have no one-dimensional analogues.

One such strategy is to difference n or more times, so that the left-hand side of the equation of lower degree becomes zero on the last application of Lemma 2.2. This enables one to exploit existing results on sums of nth powers and then to obtain further savings by realizing several variables as divisors of a fixed integer. Let $S(P, R)$ denote the number of solutions of the equation

$$x_1^n + \cdots + x_s^n = y_1^n + \cdots + y_s^n$$

with $x_i, y_i \in \mathcal{A}(P, R)$. We say that $\mu_s = \mu_s(n)$ is a permissible exponent for nth powers if one has the estimate $S(P, R) \ll P^{\mu_s + \varepsilon}$. This differs slightly from the notation of papers in which odd moments are dealt with non-trivially (see for example [11] and [39]), wherein our μ_s would be denoted by μ_{2s}.

Lemma 4.2. *Write $\varpi = 1$ when $k - n = 1$ and $\varpi = 0$ otherwise, and let μ_s denote a permissible exponent for nth powers. Then one has*

$$\int_{\mathbb{T}^2} \mathcal{F}_n^{(s)}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \ll H_n \widetilde{M}_n \left(P \widetilde{H}_n Q_n^{\lambda_s + \varepsilon} + P^\varpi Q_n^{\mu_s + \varepsilon} \right).$$

Proof. On recalling (2.1), we have by Cauchy's inequality that

$$\int_{\mathbb{T}^2} \mathcal{F}_n^{(s)}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \ll H_n V(P, R, \boldsymbol{\phi}),$$

where $V(P, R, \boldsymbol{\phi})$ denotes the number of solutions of the system

$$(4.6) \qquad \Psi_{k,n}(z_1; \mathbf{h}; \mathbf{m}) - \Psi_{k,n}(z_2; \mathbf{h}; \mathbf{m}) = \sum_{i=1}^{s}(u_i^k - v_i^k)$$

$$0 = \sum_{i=1}^{s}(u_i^n - v_i^n),$$

with $\mathbf{h}, \mathbf{m}$ as in (2.4), $1 \le z_1, z_2 \le P_n$, and $u_i, v_i \in \mathcal{A}(Q_n, R)$. We can write

$$V(P, R, \boldsymbol{\phi}) = V_0 + V_1,$$

where V_0 denotes the number of solutions with $z_1 = z_2$ and V_1 denotes the number of solutions with $z_1 \ne z_2$. Then one clearly has

$$(4.7) \qquad V_0 \ll P\widetilde{H}_n\widetilde{M}_nQ_n^{\lambda_s + \varepsilon}.$$

To count the non-diagonal solutions, we first observe that

$$\psi_{k,n}(z_1; \mathbf{h}; \mathbf{m}) - \psi_{k,n}(z_2; \mathbf{h}; \mathbf{m}) = h_1 \cdots h_n(z_1 - z_2)\Xi(\mathbf{z}; \mathbf{h}; \mathbf{m}),$$

where Ξ is a polynomial of degree $k - n - 1$ in the variables z_1 and z_2. Note also that Ξ is non-zero for all values of $\mathbf{z}$, $\mathbf{h}$, and $\mathbf{m}$ under consideration. Our strategy is to fix $\mathbf{u}$ and $\mathbf{v}$ first using the second equation in (4.6), and then fix $\mathbf{m}$ arbitrarily. The first equation in (4.6) then determines $\mathbf{h}$, $z_1 - z_2$, and Ξ to $O(P^\varepsilon)$, by a divisor estimate. Finally, when $k - n > 1$, the polynomial Ξ depends explicitly on z_1 and z_2, and it follows that z_1 and z_2 are now determined to $O(1)$. If $k - n = 1$, then we have an extra factor of P in our estimate, since the value of Ξ provides no information about z_1 and z_2. We thus have

$$V_1 \ll P^{\varpi+\varepsilon}\widetilde{M}_nQ_n^{\mu_s},$$

and the lemma follows on recalling (4.7). $\qquad\qquad\square$

We now apply Lemma 4.2 with s replaced by s_n. When the diagonal solutions dominate in (4.6), we optimize parameters in (2.10) by taking

$$P^2\widetilde{H}_{n-1}\widetilde{M}_nQ_n^{\lambda_{s_n}} \approx PH_n\widetilde{H}_n\widetilde{M}_nQ_n^{\lambda_{s_n}},$$

which reduces to $P \approx H_n^2$, or $\phi_n = (2K_n)^{-1}$. Thus we take

$$\alpha_n = (2K_n)^{-1}, \quad \beta_n = 0, \quad \text{and} \quad \gamma_n = 1$$

in (2.20), and we denote the resulting iterative process by B_n. When the non-diagonal solutions dominate, we instead take

$$P^2\widetilde{H}_{n-1}Q_n^{\lambda_{s_n}} \approx H_nP^\varpi Q_n^{\mu_{s_n}},$$

so on writing $\delta_n = \mu_{s_n} - \lambda_{s_n}$, we find after a short computation that the appropriate parameters are

$$\alpha_n = (K_n + \delta_n)^{-1}, \quad \beta_n = \delta_n - k, \quad \text{and} \quad \gamma_n = \delta_n - n + \varpi.$$

In this situation, we denote the process by B'_n to distinguish it from the diagonal case. On comparing the two terms in Lemma 4.2, we find that the diagonal solutions dominate precisely when

$$n + 1 - \delta_n - \varpi + \phi_n(\delta_n - K_n) \geq (k - \delta_n)(\phi_1 + \cdots + \phi_{n-1}).$$

When n is small relative to k, it may be useful to apply the above method with more than n differences. In this case, if $n < j < k$, a simplified version of the proof of Lemma 4.2 yields the estimate

$$\int_{\mathbb{T}^2} \mathcal{F}_j^{(s)}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \ll P^\varepsilon \widetilde{M}_j Q_j^{\mu_s}.$$

Note that since $r_{j-1} = 1$, there are no diagonal solutions in the analogue of (4.6), so we refer to the process as B'_j. We again let $\delta_j = \mu_{s_j} - \lambda_{s_j}$, and in (2.20) we take

$$\alpha_j = \delta_j^{-1}, \quad \beta_j = \delta_j - k, \quad \text{and} \quad \gamma_j = \delta_j - j.$$

When $k - n = 1$, it is sometimes more effective to difference fewer than n times and then to employ the approach used by Wooley [41] in bounding $G^*(3, 2)$. Here one is able to take a linear combination of the two equations and then exploit a version of Hua's inequality for exponential sums over arbitrary polynomials. We define the Hua exponents ν_s by taking $\nu_s = 2s - \ell$ when $1 \leq \ell \leq k$ and $s = 2^{\ell-1}$, $\nu_s = 2s - k$ when $s > 2^{k-1}$, and for all remaining $s > 2$ by linear interpolation.

Lemma 4.3. *Suppose that $j \leq n - 1$, and write $e_j = 2$ if both $k - j$ and $n - j$ are even and $j \leq n - 2$, and write $e_j = 1$ otherwise. If ν_s denotes a Hua exponent, then one has*

$$\int_{\mathbb{T}^2} \mathcal{F}_j^{(s)}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \ll H_j \widetilde{M}_j \left(P \widetilde{H}_j Q_j^{\lambda_s + \varepsilon} + P^{k+n-2j-2e_j} Q_j^{\nu_s + \varepsilon} \right).$$

Proof. In analyzing the system

$$\psi_{i,j}(z_1; \mathbf{h}; \mathbf{m}) - \psi_{i,j}(z_2; \mathbf{h}; \mathbf{m}) = D_{i,j}(\mathbf{m}) \sum_{l=1}^{s} (u_l^i - v_l^i) \qquad (i = k, n),$$

we estimate the number of solutions with $z_1 = z_2$ exactly as in the proof of Lemma 4.2. To deal with the non-diagonal solutions, we first observe that

$$\psi_{i,j}(z_1; \mathbf{h}; \mathbf{m}) - \psi_{i,j}(z_2; \mathbf{h}; \mathbf{m}) = m_1^{K_1} \cdots m_j^{K_j} h_1 \cdots h_j (z_1^{e_j} - z_2^{e_j}) \Xi_i(\mathbf{z}; \mathbf{h}; \mathbf{m}),$$

where Ξ_i is a polynomial satisfying

$$0 < |\Xi_i(\mathbf{z}; \mathbf{h}; \mathbf{m})| \ll P^{i-j-e_j+\varepsilon} \qquad (i = k, n)$$

for all values of $\mathbf{z}, \mathbf{h}$, and $\mathbf{m}$ under consideration. Here we have used the fact that the $\psi_{i,j}(z; \mathbf{h}; \mathbf{m})$ were constructed using symmetric differences, so that all powers of z occurring have the same parity. We start by fixing the

values of Ξ_k and Ξ_n in $O(P^{k+n-2j-2e_j+\varepsilon})$ ways, and we also fix $\mathbf{m}$. Then on taking a linear combination of the two equations and applying Hölder's inequality and Theorem 4 of Hua [19] as in Wooley [41], we find that the number of choices for $\mathbf{u}$ and $\mathbf{v}$ is $O(Q_j^{\nu_s+\varepsilon})$. The values of $h_1, \ldots, h_j$, z_1, and z_2 are now determined to $O(P^\varepsilon)$, and this completes the proof. $\square$

We now apply Lemma 4.3 with s replaced by s_j, following the last application of Lemma 2.2. When the underlying diagonal solutions dominate, we again take

$$\alpha_j = (2K_j)^{-1}, \quad \beta_j = 0, \quad \text{and} \quad \gamma_j = 1$$

in (2.20), and we denote the resulting process by C_j. When the non-diagonal solutions dominate, we now optimize by taking

$$P^2 \widetilde{H}_{j-1} Q_j^{\lambda_{s_j}} \approx H_j P^{k+n-2j-2e_j} Q_j^{\nu_{s_j}},$$

and on writing $\delta_j = \nu_{s_j} - \lambda_{s_j}$, a brief computation shows that this corresponds to the choice

$$\alpha_j = (K_j + \delta_j)^{-1}, \quad \beta_j = \delta_j - k, \quad \text{and} \quad \gamma_j = \delta_j + k + n - 3j - 2e_j.$$

In this case, we denote the iterative process by C_j'. An examination of the terms in Lemma 4.3 reveals that the diagonal solutions dominate whenever

$$3j + 2e_j + 1 - k - n - \delta_j + \phi_j(\delta_j - K_j) \geq (k - \delta_j)(\phi_1 + \cdots + \phi_{j-1}).$$

It is in some ways natural to view the ad-hoc analysis described in Lemmas 4.2 and 4.3 as a type of Hardy-Littlewood dissection, with the non-diagonal and diagonal solutions corresponding respectively to the major and minor arcs (see for example [33]). An actual Hardy-Littlewood dissection is described in the next section.

5. Major and Minor Arc Estimates

In the final stages of the iteration, it is often useful to estimate the integral in (2.10) by means of a Hardy-Littlewood dissection. Throughout this section, we assume that $K_j = k$, so we may work with the polynomials polynomials $\Psi_{k,j}$ and $\Psi_{n,j}$. From Vaughan and Wooley [32], we have

$$(5.1) \qquad \Psi_{i,j}(z; \mathbf{h}; \mathbf{m}) = \sum_{u,\mathbf{v}} \frac{i! \, 2^j \widetilde{h} \widetilde{m}^{k-i} z^u (h_1 m_1^k)^{2v_1} \cdots (h_j m_j^k)^{2v_j}}{u!(2v_1+1)! \cdots (2v_j+1)!},$$

where the summation is over non-negative integers $u, v_1, \ldots, v_j$ satisfying

$$u + 2v_1 + \cdots + 2v_j = i - j,$$

and where we have written $\widetilde{h} = h_1 \cdots h_j$ and $\widetilde{m} = m_1 \cdots m_j$. We shall find it convenient to introduce the notation

$$\tau_j = 2^{1+j-k} \quad \text{and} \quad J = 2^{n-j-1}.$$

The processes described in this section will be applicable only when the conditions

(5.2) $$1 \leq j \leq n - 2,$$

(5.3) $$2\tau_j(k - j) \leq 1,$$

(5.4) $$(k - j)(k - n)\tau_j \leq (k - n + 1)(1 - J\tau_j),$$

(5.5) $$(k - n - 1)(\phi_1 + \cdots + \phi_j) \leq 1 - \tau_j(J + k - j),$$

(5.6) $$(k - n - 1)(\phi_1 + \cdots + \phi_j) \leq 1 - J\tau_j(k - n + 1)$$

are all satisfied. We start by showing that, under these conditions, an unusually large value of the exponential sum $F_j(\boldsymbol{\alpha})$ defined by (2.2) implies that $\boldsymbol{\alpha}$ lies within a standard set of major arcs.

Lemma 5.1. *Suppose that $F_j(\boldsymbol{\alpha}) \geq \widetilde{M}_j H_j \widetilde{H}_j P^{2-2\tau_j+\varepsilon}$ for some $\varepsilon > 0$ and that conditions* (5.2)–(5.6) *are satisfied. Then there exists a positive number δ and integers q, a_k, and a_n with $(q, a_k, a_n) = 1$ such that*

$$1 \leq q \leq P, \quad |q\alpha_k - a_k| \leq P^{1-\delta}Q_j^{-k}, \quad \text{and} \quad |q\alpha_n - a_n| \leq P^{1-\delta}Q_j^{-n},$$

whenever P is sufficiently large.

Proof. First of all, we have by Cauchy's inequality that

(5.7) $$F_j(\boldsymbol{\alpha}) \leq H_j \sum_{m_1,\ldots,m_j} \sum_{h_1,\ldots,h_j} |G_j(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m})|^2,$$

where G_j is as in (2.3) and where the summations range over $\mathbf{m}$ and $\mathbf{h}$ in the intervals (2.4). Therefore, if $F_j(\boldsymbol{\alpha}) \geq \widetilde{M}_j H_j \widetilde{H}_j P^{2-2\tau_j+\varepsilon}$, then there exist U and A such that

(5.8) $$\sum_{(\mathbf{h},\mathbf{m})\in\mathcal{A}} |G_j(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m})|^2 \gg \widetilde{M}_j \widetilde{H}_j P^{2-2\tau_j+\varepsilon}(\log P)^{-2},$$

where we have written

$$\mathcal{A} = \{(\mathbf{h}, \mathbf{m}) : U < h_1 \cdots h_j \leq 2U \quad \text{and} \quad A \leq |G_j(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m})| \leq 2A\}.$$

It follows easily from (5.8) that

(5.9) $$A \gg P^{1-\tau_j+\varepsilon} \quad \text{and} \quad \operatorname{card}(\mathcal{A}) \gg \widetilde{M}_j \widetilde{H}_j P^{2-2\tau_j+\varepsilon} A^{-2},$$

where the value of ε occurring in these two bounds may be smaller than that in (5.8); see the discussion at the end of §1. In the remainder of the

proof, it is to be understood that each new ε that appears is to be taken sufficiently small in terms of any previous ε.

Choose any $(\mathbf{h}, \mathbf{m}) \in \mathcal{A}$. Since $|G_j(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m})| \geq A$, we may apply Theorem 5.1 of Baker [2] (see also [1], [4]) to obtain integers w and b with

$$(5.10) \qquad 1 \leq w \leq P^{k-j+\varepsilon} A^{j-k} \qquad \text{and} \qquad |C_k \alpha_k w - b| < P^\varepsilon A^{j-k},$$

where we have written

$$(5.11) \qquad C_i = \frac{i! \, 2^j}{(i-j)!} h_1 \cdots h_j (m_1 \cdots m_j)^{k-i}$$

for the coefficient of z^{i-j} in the polynomial $\Psi_{i,j}$. Clearly, we may further assume that $(w, b) = 1$. We now let

$$x = w(h_1 \cdots h_j) \qquad \text{and} \qquad \theta = \frac{k!}{(k-j)!} 2^j \alpha_k,$$

and write $z = z(x)$ for the nearest integer to $x\theta$. Then by (5.10) one has $1 \leq x \leq X$, where $X \ll U P^{k-j+\varepsilon} A^{j-k}$ and $|x\theta - z| < \zeta$, where $\zeta \ll P^\varepsilon A^{j-k}$. Moreover, as $(\mathbf{h}, \mathbf{m})$ runs over the set $\mathcal{A}$, one sees from (5.9) and a divisor estimate that x assumes N distinct values, where

$$(5.12) \qquad N \gg \widetilde{H}_j P^{2-2\tau_j+\varepsilon} A^{-2}.$$

In view of condition (5.2), we have $N > 24X\zeta$ for P sufficiently large, and we may therefore apply Lemma 14 of Birch and Davenport [6] (see also Lemma 5.2 of Baker [2]) to obtain fixed coprime integers u and v with the property that $u/v = z/x$ for all x under consideration. It follows that $v|x$ for each x, and hence $N \ll U P^{k-j+\varepsilon} A^{j-k} v^{-1}$. We therefore obtain from (5.12) that

$$v \ll U \widetilde{H}_j^{-1} P^{k-j-2+2\tau_j+\varepsilon} A^{2+j-k},$$

and, since $x \gg U$, it follows that

$$|v\theta - u| = vx^{-1}|x\theta - z| \ll \widetilde{H}_j^{-1} P^{k-j-2+2\tau_j+\varepsilon} A^{2(1+j-k)}.$$

On setting $r = 2^j k(k-1)\cdots(k-j+1)v$, we then see that

$$(5.13) \qquad r \ll P^{k-j-2+2\tau_j+\varepsilon} A^{2+j-k} \ll P^{(k-j)\tau_j-\varepsilon}$$

and

$$(5.14) \qquad |r\alpha_k - u| \ll \widetilde{H}_j^{-1} P^{k-j-2+2\tau_j+\varepsilon} A^{2(1+j-k)} \ll \widetilde{H}_j^{-1} P^{(k-j)(2\tau_j-1)-\varepsilon},$$

on noting that the ε occurring in (5.10) can be chosen sufficiently small in terms of the one occurring in (5.9). We may further assume that $(r, u) = 1$, since (5.13) and (5.14) are preserved when dividing out common factors. We note here that r and u depend on the set $\mathcal{A}$, but not on any particular choice of $\mathbf{h}$ or $\mathbf{m}$.

Our strategy now is to difference the sum G_j until α_n occurs only in the linear term, at which point a rational approximation can be obtained from (5.14) via Baker's final coefficient lemma. Define the ordinary differencing operator Δ_l by

$$\Delta_l(z; \mathbf{g}) = \Delta_{l,\mathbf{1}}^*(z; \mathbf{g}; \mathbf{1}),$$

where $\Delta_{l,\mathbf{K}}^*$ is as in §2, and write

$$\widetilde{\Psi}_{i,j}(z; \mathbf{h}; \mathbf{m}; \mathbf{g}) = \Delta_{n-j-1}(\Psi_{i,j}(z; \mathbf{h}; \mathbf{m}); \mathbf{g}).$$

By applying Hölder's inequality in (5.8), we obtain

$$\widetilde{M}_j \widetilde{H}_j P^{J(1-\tau_j)+\varepsilon} \ll \sum_{(\mathbf{h},\mathbf{m}) \in \mathcal{A}} |G_j(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m})|^J.$$

Then by Weyl differencing (see for example [2], Lemma 3.8), we find that

(5.15) $$\widetilde{M}_j \widetilde{H}_j P^{J(1-\tau_j)+\varepsilon} \ll P^{J-n+j} \sum_{(\mathbf{h},\mathbf{m}) \in \mathcal{A}} \sum_{\mathbf{g}} |S(\mathbf{h}, \mathbf{m}, \mathbf{g})|,$$

where the second summation is over $g_{j+1}, \ldots, g_{n-1} \in [1, P_j]$, and where

$$S(\mathbf{h}, \mathbf{m}, \mathbf{g}) = \sum_{z \in \mathcal{I}(\mathbf{g})} e(\alpha_k \widetilde{\Psi}_{k,j}(z; \mathbf{h}; \mathbf{m}; \mathbf{g}) + \alpha_n \widetilde{\Psi}_{n,j}(z; \mathbf{h}; \mathbf{m}; \mathbf{g}))$$

for some interval $\mathcal{I}(\mathbf{g}) \subseteq [1, P_j]$. By (5.15), there exist V, B, and D such that

(5.16) $$\widetilde{M}_j \widetilde{H}_j P^{n-j-J\tau_j+\varepsilon} \ll \sum_{(\mathbf{h},\mathbf{m},\mathbf{g}) \in \mathcal{B}} |S(\mathbf{h}, \mathbf{m}, \mathbf{g})|,$$

where $\mathcal{B}$ denotes the set of all $(\mathbf{h}, \mathbf{m}, \mathbf{g})$ for which

$$(\mathbf{h}, \mathbf{m}) \in \mathcal{A}, \qquad V < g_{j+1} \cdots g_{n-1} \leq 2V,$$

$$B < |S(\mathbf{h}, \mathbf{m}, \mathbf{g})| \leq 2B, \qquad \text{and} \qquad (r, h_1 \cdots h_j g_{j+1} \cdots g_{n-1}) = D.$$

Note that by (5.16) we have

(5.17) $$B \gg P^{n-j-J\tau_j+\varepsilon} V^{-1} \gg P^{1-J\tau_j+\varepsilon}$$

and

(5.18) $$\mathrm{card}(\mathcal{B}) \gg \widetilde{M}_j \widetilde{H}_j P^{n-j-J\tau_j+\varepsilon} B^{-1}.$$

We now choose $(\mathbf{h}, \mathbf{m}, \mathbf{g}) \in \mathcal{B}$, and write

$$\widetilde{h} = h_1 \cdots h_j, \qquad \widetilde{m} = m_1 \cdots m_j, \qquad \widetilde{g} = g_{j+1} \cdots g_{n-1}.$$

Further, write $\mathbf{hm}^k$ for the vector $(h_1 m_1^k, \ldots, h_j m_j^k)$, and define $p_i(\mathbf{h}, \mathbf{m}, \mathbf{g})$ by

$$\alpha_k \widetilde{\Psi}_{k,j}(z; \mathbf{h}; \mathbf{m}; \mathbf{g}) + \alpha_n \widetilde{\Psi}_{n,j}(z; \mathbf{h}; \mathbf{m}; \mathbf{g}) = \sum_{i=0}^{k-n+1} p_i(\mathbf{h}, \mathbf{m}, \mathbf{g}) z^i.$$

It suffices for our purposes to note that

$$p_i(\mathbf{h}, \mathbf{m}, \mathbf{g}) = \widetilde{h}\widetilde{g}\,\Phi_i(\mathbf{hm}^k, \mathbf{g})\alpha_k \qquad (2 \le i \le k - n + 1)$$

and

$$p_1(\mathbf{h}, \mathbf{m}, \mathbf{g}) = \widetilde{h}\widetilde{g}\left(\Phi_1(\mathbf{hm}^k, \mathbf{g})\alpha_k + 2^j n!\widetilde{m}^{k-n}\alpha_n\right),$$

where $\Phi_i(\mathbf{x}, \mathbf{y})$ is a homogeneous polynomial of degree $k - n - i + 1$ in $n - 1$ variables, with integer coefficients. Since

$$\Phi_i(\mathbf{hm}^k, \mathbf{g}) \ll P^{k-n-i+1+\varepsilon},$$

we have by (5.14) that

$$(5.19) \qquad |\widetilde{h}\widetilde{g}\,\Phi_i(\mathbf{hm}^k, \mathbf{g})(r\alpha_k - u)| \ll P^{2\tau_j(k-j)-i-\varepsilon}.$$

Moreover, $|S(\mathbf{h}, \mathbf{m}, \mathbf{g})| \ge B$, so on recalling (5.3) and (5.4), it follows with a little computation from (5.13), (5.17), and (5.19) that we may apply [2], Lemma 4.6, to obtain a rational approximation to $p_1(\mathbf{h}, \mathbf{m}, \mathbf{g})$. On writing

$$d = (r, u\widetilde{h}\widetilde{g}\,\Phi_2(\mathbf{hm}^k, \mathbf{g}), \ldots, u\widetilde{h}\widetilde{g}\,\Phi_{k-n+1}(\mathbf{hm}^k, \mathbf{g})),$$

we find that there exists a natural number $t \le 2(k - n + 1)^2$ such that

$$trd^{-1} \le P^{k-n+1+\varepsilon}B^{n-k-1},$$

$$(5.20) \qquad t|\widetilde{h}\widetilde{g}\,\Phi_i(\mathbf{hm}^k, \mathbf{g})(r\alpha_k - u)|d^{-1} \le P^{k-n+1-i+\varepsilon}B^{n-k-1}$$

whenever $2 \le i \le k - n + 1$, and

$$(5.21) \qquad ||trd^{-1}p_1(\mathbf{h}, \mathbf{m}, \mathbf{g})|| \le P^{k-n+\varepsilon}B^{n-k-1}.$$

It is clear that $D|d$, and since $(r, u) = 1$ and $\Phi_{k-n+1} \ll 1$, we also have $d \ll D$.

Now write $x' = t\widetilde{h}\widetilde{g}\widetilde{m}^{k-n}D^{-1}$ and $\theta' = 2^j n!r\alpha_n$, and let z' be the nearest integer to $x'\theta'$. Then one sees easily by applying the triangle inequality and using (5.20) and (5.21) that

$$(5.22) \qquad |x'\theta' - z'| \ll P^{k-n+\varepsilon}B^{n-k-1},$$

and we also note that $x' \ll X' = UV\widetilde{M}_j^{k-n}D^{-1}$. As $(\mathbf{h}, \mathbf{m}, \mathbf{g})$ runs through $\mathcal{B}$, we see from (5.18) that there are

$$N' \gg \widetilde{M}_j\widetilde{H}_j P^{n-j-J\tau_j+\varepsilon}B^{-1}$$

distinct values of x', whence it follows from (5.6), after some computation, that we may once again apply [6], Lemma 14. Thus we obtain coprime integers u' and v' such that $u'/v' = z'/x'$ for all x'. It follows that $v'|x'$, whence $N' \ll X'/v'$ and

$$(5.23) \qquad v' \ll \widetilde{M}_j^{k-n-1}\widetilde{H}_j^{-1}P^{J\tau_j-n+j}UVBD^{-1}.$$

Since $x' \gg UV\widetilde{M}_j^{k-n}D^{-1}$, we have from (5.22) that

$$(5.24) \qquad |v'\theta' - u'| = \frac{v'}{x'}|x'\theta' - z'| \ll (\widetilde{M}_j\widetilde{H}_j)^{-1}P^{k-2n+j+J\tau_j+\varepsilon}B^{n-k}.$$

We now put

$$q = 2^j n! v'r, \qquad a_n = u', \quad \text{and} \qquad a_k = 2^j n! v'u.$$

In view of (5.5), we can deduce from (5.13) and (5.23) that $q \leq P^{1-\delta}$ for some $\delta > 0$. On recalling (5.6) and (5.17), some mild computation reveals that the bound $|q\alpha_n - a_n| \leq P^{1-\delta}Q_j^{-n}$ follows from (5.24) and that the inequality $|q\alpha_k - a_k| \leq P^{1-\delta}Q_j^{-k}$ is a consequence of (5.20) and (5.23). Finally, one may divide out common factors in the above inequalities to ensure that $(q, a_k, a_n) = 1$, and this completes the proof. $\qquad \square$

We now define a set of major arcs $\mathfrak{M}_j$ to be the union of the rectangles

$$(5.25) \qquad \mathfrak{M}_j(q, \mathbf{a}) = \{\boldsymbol{\alpha} \in \mathbb{T}^2 : |q\alpha_i - a_i| \leq PR^{k(j-k)}Q_j^{-i} \ (i = k, n)\}$$

with $0 \leq a_k, a_n \leq q \leq P$ and $(q, a_k, a_n) = 1$. Further, write $\mathfrak{m}_j = \mathbb{T}^2 \setminus \mathfrak{M}_j$ for the minor arcs. Then Lemma 5.1 tells us that

$$(5.26) \qquad \sup_{\boldsymbol{\alpha} \in \mathfrak{m}_j} |F_j(\boldsymbol{\alpha})| \ll \widetilde{M}_j H_j \widetilde{H}_j P^{2-2\tau_j+\varepsilon}.$$

Our next task is to obtain suitable estimates for $F_j(\boldsymbol{\alpha})$ on the major arcs. The $j = 1$ case of the following lemma appears in an unpublished manuscript of Wooley (c. 1990), and the general case follows the same pattern of proof (see also Vaughan and Wooley [32], Lemma 4.7).

Lemma 5.2. *Suppose that $k > n \geq 3$ and $j \leq k - 2$, and write*

$$\varpi_j = \max\left(0, 2(k - n)(k - j)^{-1} - 1\right).$$

Then whenever $\boldsymbol{\alpha} \in \mathfrak{M}_j(q, \mathbf{a}) \subseteq \mathfrak{M}_j$, one has

$$F_j(\boldsymbol{\alpha}) \ll P^{2+\varepsilon}q^{\varpi_j}\widetilde{M}_j H_j \widetilde{H}_j(q + |q\alpha_n - a_n|Q_j^n + |q\alpha_k - a_k|Q_j^k)^{-2/(k-j)}.$$

Proof. Suppose that $\boldsymbol{\alpha} \in \mathfrak{M}_j(q, \mathbf{a}) \subseteq \mathfrak{M}_j$, and write $\beta_i = \alpha_i - a_i/q$ for $i = k, n$. We first observe that, for any given $\mathbf{h}$ and $\mathbf{m}$, one has

$$(5.27) \qquad G_j(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m}) = q^{-1} \sum_{-q/2 < b \leq q/2} \sigma(q, \mathbf{a}, b, \mathbf{h}, \mathbf{m})T(\boldsymbol{\beta}, b, \mathbf{h}, \mathbf{m}),$$

where

$$\sigma(q, \mathbf{a}, b, \mathbf{h}, \mathbf{m}) = \sum_{r=1}^{q} e\big((a_n \Psi_{n,j}(r; \mathbf{h}; \mathbf{m}) + a_k \Psi_{k,j}(r; \mathbf{h}; \mathbf{m}) + br)/q\big)$$

and
$$T(\boldsymbol{\beta}, b, \mathbf{h}, \mathbf{m}) = \sum_{1 \le z \le P_j} e(\beta_n \Psi_{n,j}(z; \mathbf{h}; \mathbf{m}) + \beta_k \Psi_{k,j}(z; \mathbf{h}; \mathbf{m}) - bz/q).$$

For a fixed $\mathbf{h}$ and $\mathbf{m}$, define the integers $p_0, \ldots, p_{k-j}$ by
$$a_n \Psi_{n,j}(r; \mathbf{h}; \mathbf{m}) + a_k \Psi_{k,j}(r; \mathbf{h}; \mathbf{m}) = \sum_{i=0}^{k-j} p_i r^i.$$

Then by (5.1), one has in particular that
$$p_{k-j} = C_k a_k, \quad p_{n-j} = C_n a_n + D_{k,n} a_k, \quad \text{and} \quad p_1 = D_{n,j+1} a_n + D_{k,j+1} a_k,$$
where C_i is as in (5.11) and where

$$(5.28) \qquad D_{i,l} = \frac{C_i (i-j)!}{(l-j)!} \sum_{\substack{v_1, \ldots, v_j \ge 0 \\ 2v_1 + \cdots + 2v_j = i-l}} \frac{(h_1 m_1^k)^{2v_1} \cdots (h_j m_j^k)^{2v_j}}{(2v_1 + 1)! \cdots (2v_j + 1)!}$$

is the coefficient of z^{l-j} in $\Psi_{i,j}(z; \mathbf{h}; \mathbf{m})$. It follows that
$$(q, p_{k-j}, \ldots, p_2, p_1 + b) \ll (q, \widetilde{h} a_k, \widetilde{h} \widetilde{m}^{k-n} a_n, b) \ll d(q, \widetilde{h}, b),$$
where we have written
$$\widetilde{h} = h_1 \cdots h_j, \quad \widetilde{m} = m_1 \cdots m_j, \quad \text{and} \quad d = (q, a_k, \widetilde{m}^{k-n} a_n),$$
and we therefore deduce from Lemma 4.2 of Baker [2] (see also Theorem 7.1 of Vaughan [30]) that
$$(5.29) \qquad \sigma(q, \mathbf{a}, b, \mathbf{h}, \mathbf{m}) \ll q^{(k-j-1)/(k-j)+\varepsilon}[d(q, \widetilde{h}, b)]^{1/(k-j)}.$$

Now for a fixed b, write
$$\Phi(\gamma) = \beta_n \Psi_{n,j}(\gamma; \mathbf{h}; \mathbf{m}) + \beta_k \Psi_{k,j}(\gamma; \mathbf{h}; \mathbf{m}) - b\gamma/q.$$

Then, as in the proof of Vaughan and Wooley [32], Lemma 4.7, one easily verifies that
$$\frac{b}{q} + \Phi'(\gamma) = \frac{n!}{(n-j-1)!} \beta_n \widetilde{m}^{-n} J_n + \frac{k!}{(k-j-1)!} \beta_k \widetilde{m}^{-k} J_k,$$
where
$$J_i = \int_{\gamma - h_1 m_1^k}^{\gamma + h_1 m_1^k} \int_{y_1 - h_2 m_2^k}^{y_1 + h_2 m_2^k} \cdots \int_{y_{j-1} - h_j m_j^k}^{y_{j-1} + h_j m_j^k} y_j^{i-j-1} dy_j \cdots dy_1.$$

Thus when $|\gamma| \le P_j$ and $\boldsymbol{\alpha} \in \mathfrak{M}_j(q, \mathbf{a})$, one finds on recalling (5.25) that
$$\left| \frac{b}{q} + \Phi'(\gamma) \right| \ll \widetilde{h} \widetilde{m}^{k-n} (PR^k)^{n-j-1} |\beta_n| + \widetilde{h} (PR^k)^{k-j-1} |\beta_k| \ll (qR)^{-1}.$$

Since R is a positive power of P, we may conclude that

$$(5.30) \qquad \left| \frac{b}{q} + \Phi'(\gamma) \right| \leq \frac{1}{4q}$$

when P is sufficiently large, and hence when $|b| \leq q/2$ one has

$$(5.31) \qquad |\Phi'(\gamma)| \leq \frac{1}{4q} + \frac{|b|}{q} \leq \frac{3}{4}.$$

We may therefore apply Vaughan [30], Lemma 4.2, to conclude that

$$(5.32) \qquad T(\boldsymbol{\beta}, b, \mathbf{h}, \mathbf{m}) = \sum_{u=-1}^{1} I(\boldsymbol{\beta}, b, \mathbf{h}, \mathbf{m}, u) + O(1),$$

where

$$I(\boldsymbol{\beta}, b, \mathbf{h}, \mathbf{m}, u) = \int_0^{P_j} e(\Phi(\gamma) - \gamma u) \, d\gamma.$$

In light of (5.31), we may integrate by parts to obtain

$$I(\boldsymbol{\beta}, b, \mathbf{h}, \mathbf{m}, \pm 1) \ll 1,$$

and so it follows from (5.32) that

$$(5.33) \qquad T(\boldsymbol{\beta}, b, \mathbf{h}, \mathbf{m}) = I(\boldsymbol{\beta}, b, \mathbf{h}, \mathbf{m}, 0) + O(1).$$

When $b \neq 0$, we find using (5.30) that $|\Phi'(\gamma)| \geq |b|/2q$, so another integration by parts gives

$$I(\boldsymbol{\beta}, b, \mathbf{h}, \mathbf{m}, 0) \ll \frac{q}{|b|},$$

and hence by (5.33) we obtain

$$T(\boldsymbol{\beta}, b, \mathbf{h}, \mathbf{m}) \ll \frac{q}{|b|}.$$

Thus we deduce from (5.27), (5.29), and (5.33) that

$$(5.34) \qquad G_j(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m}) = q^{-1} \sigma(q, \mathbf{a}, 0, \mathbf{h}, \mathbf{m}) I(\boldsymbol{\beta}, 0, \mathbf{h}, \mathbf{m}, 0) + E,$$

where

$$E \ll d^{1/(k-j)} q^{(k-j-1)/(k-j)+\varepsilon} \sum_{1 \leq b \leq q/2} b^{-1}(q, b)^{1/(k-j)}.$$

On using a standard estimate for the divisor function, one finds that

$$\sum_{1 \leq b \leq q/2} b^{-1}(q, b)^{1/(k-j)} \ll \sum_{r|q} r^{1/(k-j)} \sum_{1 \leq t \leq q/2r} (tr)^{-1} \ll q^{\varepsilon},$$

and it now follows from (5.29) and (5.34) that

$$(5.35) \quad G_j(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m}) \ll d^{1/(k-j)} q^{\varepsilon - 1/(k-j)} \left((q, \widetilde{h})^{1/(k-j)} |I_0(\boldsymbol{\beta}, \mathbf{h}, \mathbf{m})| + q \right),$$

where we have written $I_0(\boldsymbol{\beta}, \mathbf{h}, \mathbf{m})$ for $I(\boldsymbol{\beta}, 0, \mathbf{h}, \mathbf{m}, 0)$. By Theorem 7.3 of Vaughan [30], we have

$$(5.36) \qquad I_0(\boldsymbol{\beta}, \mathbf{h}, \mathbf{m}) \ll P(1 + |\eta_1|P + \cdots + |\eta_{k-j}|P^{k-j})^{-1/(k-j)},$$

where the real numbers η_i are defined by

$$\beta_n \Psi_{n,j}(\gamma; \mathbf{h}; \mathbf{m}) + \beta_k \Psi_{k,j}(\gamma; \mathbf{h}; \mathbf{m}) = \sum_{i=0}^{k-j} \eta_i \gamma^i.$$

One has in particular that

$$\eta_{k-j} = C_k \beta_k \quad \text{and} \quad \eta_{n-j} = C_n \beta_n + D_{k,n} \beta_k,$$

where C_i and $D_{i,l}$ are as in (5.11) and (5.28). It follows after some computation that

$$P^{n-j}|\eta_{n-j}| + P^{k-j}|\eta_{k-j}| \gg \widetilde{h}\widetilde{m}^{k-n}P^{n-j}R^{jk(n-k)}|\beta_n| + \widetilde{h}P^{k-j}|\beta_k|,$$

and we therefore deduce from (5.36) that

$$(5.37) \qquad I_0(\boldsymbol{\beta}, \mathbf{h}, \mathbf{m}) \ll PR^{jk}(1 + \widetilde{h}\widetilde{m}^{k-n}P^{n-j}|\beta_n| + \widetilde{h}P^{k-j}|\beta_k|)^{-1/(k-j)}.$$

We are now ready to sum over $\mathbf{h}$ and $\mathbf{m}$ lying in the ranges (2.4). By (5.35), we have

$$\sum_{\mathbf{h}} |G_j(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m})|^2 \ll q^{\varepsilon}(d/q)^{2/(k-j)} \left(\sum_{\mathbf{h}} (q, \widetilde{h})^{2/(k-j)} |I_0(\boldsymbol{\beta}, \mathbf{h}, \mathbf{m})|^2 + \widetilde{H}_j q^2 \right).$$

Now let $\ell = k$ or n be the index for which the quantity $\widetilde{h}\widetilde{m}^{k-\ell}P^{\ell-j}|\beta_\ell|$ is minimal. Then one sees from (5.37) that

$$\sum_{\mathbf{h}} (q, \widetilde{h})^{2/(k-j)} |I_0(\boldsymbol{\beta}, \mathbf{h}, \mathbf{m})|^2$$

$$\ll P^{2+\varepsilon} \sum_{e|q} e^{2/(k-j)} \sum_{g \le \widetilde{H}_j/e} \min\left\{ 1, \left(eg\widetilde{m}^{k-\ell}P^{\ell-j}|\beta_\ell| \right)^{-2/(k-j)} \right\}$$

$$\ll P^{2+\varepsilon} \widetilde{H}_j \min\left\{ 1, \left(P^{\ell-j}\widetilde{H}_j\widetilde{M}_j^{k-\ell}|\beta_\ell| \right)^{-2/(k-j)} \right\}.$$

A short calculation reveals that

$$P^{\ell-j}\widetilde{H}_j\widetilde{M}_j^{k-\ell} \approx Q_j^\ell,$$

and it follows that

$$(5.38) \qquad \sum_{\mathbf{h}} |G_j(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m})|^2 \ll d^{2/(k-j)}\widetilde{H}_j \left(P^{2+\varepsilon}\Delta(\boldsymbol{\alpha}) + q^{2-2/(k-j)+\varepsilon} \right),$$

where we write

$$(5.39) \qquad \Delta(\boldsymbol{\alpha}) = (q + |q\beta_n|Q_j^n + |q\beta_k|Q_j^k)^{-2/(k-j)}$$

whenever $\boldsymbol{\alpha} \in \mathfrak{M}_j(q, \mathbf{a})$. Moreover, it is easy to see from (5.25) that the first term in (5.38) is always larger than the second. Therefore, we have by Cauchy's inequality that

$$F_j(\boldsymbol{\alpha}) \leq H_j \sum_{\mathbf{m}} \sum_{\mathbf{h}} |G_j(\boldsymbol{\alpha}; \mathbf{h}, \mathbf{m})|^2$$

(5.40)
$$\ll H_j \widetilde{H}_j P^{2+\varepsilon} \Delta(\boldsymbol{\alpha}) \sum_{\mathbf{m}} (q, a_k, \widetilde{m}^{k-n} a_n)^{2/(k-j)}.$$

But since $(q, a_k, a_n) = 1$, one has

$$(q, a_k, \widetilde{m}^{k-n} a_n) \leq (q, a_k, a_n)(q, \widetilde{m}^{k-n}) = (q, \widetilde{m}^{k-n}) \leq (q, \widetilde{m})^{k-n},$$

and hence

$$\sum_{\mathbf{m}} (q, a_k, \widetilde{m}^{k-n} a_n)^{2/(k-j)} \ll \widetilde{M}_j P^\varepsilon \sum_{e|q} e^{2(k-n)/(k-j)-1} \ll P^\varepsilon \widetilde{M}_j q^{\varpi_j}.$$

The lemma now follows immediately from (5.39) and (5.40). □

Before proceeding, we observe that one has $\varpi_j = 0$ whenever $j \leq 2n - k$, and this will often be the case in our applications, in light of conditions (5.2)–(5.6). We next obtain a mean value estimate that will be used to bound the major arc contribution to the integral in (2.10). The technique of the proof is similar to that of Brüdern [8], Lemma 2.

Lemma 5.3. *Suppose that $\phi_1 + \cdots + \phi_j \leq 1/2$. Then one has*

$$\int_{\mathfrak{M}_j} F_j(\boldsymbol{\alpha})^{k-j} |f_j(\boldsymbol{\alpha})|^4 d\boldsymbol{\alpha} \ll P^{\varpi_j(k-j)+\varepsilon} F_j(0)^{k-j} Q_j^{4-k-n},$$

where ϖ_j is as in the statement of Lemma 5.2.

Proof. Write $\mathcal{J}$ for the integral in the statement of the lemma. When $\boldsymbol{\alpha} \in \mathfrak{M}_j(q, \mathbf{a}) \subseteq \mathfrak{M}_j$, we have from Lemma 5.2 that

$$F_j(\boldsymbol{\alpha}) \ll P^{2+\varepsilon} q^{\varpi_j} \widetilde{M}_j H_j \widetilde{H}_j (q + Q_j^k |q\beta_k|)^{-1/(k-j)} (q + Q_j^n |q\beta_n|)^{-1/(k-j)},$$

where we have written $\beta_i = \alpha_i - a_i/q$ for $i = k, n$, and it follows that

$$\mathcal{J} \ll F_j(0)^{k-j+\varepsilon} \sum_{\substack{1 \leq q \leq P \\ \mathbf{a} \in [0,q]^2}} q^{\varpi_j(k-j)-2} \int_{\mathfrak{M}_j(q, \mathbf{a})} \frac{|f_j(\boldsymbol{\alpha})|^4 \, d\boldsymbol{\alpha}}{(1 + Q_j^k |\beta_k|)(1 + Q_j^n |\beta_n|)}.$$

On making a change of variables and expanding the fourth power of $|f_j(\boldsymbol{\alpha})|$, we find that

$$\mathcal{J} \ll F_j(0)^{k-j+\varepsilon} \sum_{\substack{1 \leq q \leq P \\ \mathbf{a} \in [1,q]^2}} q^{\varpi_j(k-j)-2} \int_{\mathcal{B}_k(q) \times \mathcal{B}_n(q)} \frac{S(q, \mathbf{a}, \boldsymbol{\beta}) \, d\boldsymbol{\beta}}{(1 + Q_j^k |\beta_k|)(1 + Q_j^n |\beta_n|)},$$

where we have written

$$S(q, \mathbf{a}, \boldsymbol{\beta}) = \sum_{\mathbf{x} \in \mathcal{A}(Q_j, R)^4} e\left(\left(\frac{a_k}{q} + \beta_k\right) S_k(\mathbf{x})\right) e\left(\left(\frac{a_n}{q} + \beta_n\right) S_n(\mathbf{x})\right),$$

$$S_i(\mathbf{x}) = x_1^i + x_2^i - x_3^i - x_4^i,$$

and

$$\mathcal{B}_i(q) = [-q^{-1} P R^{k(j-k)} Q_j^{-i}, q^{-1} P R^{k(j-k)} Q_j^{-i}].$$

After bringing the summation over $\mathbf{x}$ and the factors of $e(a_i S_i(\mathbf{x})/q)$ outside the integral and noting that $|e(\beta_i S_i(\mathbf{x}))| = 1$, we obtain the estimate

$$(5.41) \qquad \int_{\mathfrak{M}_j} F_j(\boldsymbol{\alpha})^{k-j} |f_j(\boldsymbol{\alpha})|^4 d\boldsymbol{\alpha} \ll F_j(0)^{k-j+\varepsilon} \mathcal{S}(P) \mathcal{I}(P),$$

where

$$(5.42) \quad \mathcal{S}(P) = \sum_{1 \leq q \leq P} q^{\varpi_j(k-j)-2} \sum_{\mathbf{x} \in \mathcal{A}(Q_j, R)^4} \left| \sum_{\mathbf{a} \in [1,q]^2} e\left(\frac{a_k S_k(\mathbf{x}) + a_n S_n(\mathbf{x})}{q}\right) \right|$$

and

$$(5.43) \qquad \mathcal{I}(P) = \int_{-PQ_j^{-k}}^{PQ_j^{-k}} \int_{-PQ_j^{-n}}^{PQ_j^{-n}} (1 + Q_j^k |\beta_k|)^{-1} (1 + Q_j^n |\beta_n|)^{-1} d\beta_n \, d\beta_k.$$

The summation over $\mathbf{a}$ in (5.42) is q^2 if

$$x_1^k + x_2^k - x_3^k - x_4^k \equiv x_1^n + x_2^n - x_3^n - x_4^n \equiv 0 \pmod{q}$$

and zero otherwise. By Hua's Lemma (for example), the number of q and $\mathbf{x}$ for which

$$x_1^k + x_2^k - x_3^k - x_4^k = x_1^n + x_2^n - x_3^n - x_4^n = 0$$

is at most $PQ_j^{2+\varepsilon}$, while a divisor estimate shows that there are at most $P^\varepsilon Q_j^4$ remaining q and $\mathbf{x}$ that give a non-zero contribution to (5.42). One therefore has

$$\mathcal{S}(P) \ll P^{\varpi_j(k-j)} \left(PQ_j^{2+\varepsilon} + P^\varepsilon Q_j^4\right) \ll P^{\varpi_j(k-j)+\varepsilon} Q_j^4,$$

on noting that the hypothesis in the statement of the lemma implies that $P \leq Q_j^2$. Finally, on making a change of variables in (5.43), one finds that

$$\mathcal{I}(P) \ll Q_j^{-k-n} \int_{-P}^{P} (1 + |\beta_k|)^{-1} d\beta_k \int_{-P}^{P} (1 + |\beta_n|)^{-1} d\beta_n \ll Q_j^{-k-n} (\log P)^2,$$

and the lemma now follows from (5.41). $\qquad \square$

In certain situations, it may be advantageous to prove a similar lemma with the fourth power of f_j replaced by some higher power, so that the assumption on $\phi_1 + \cdots + \phi_j$ may be weakened. However, because of the other restrictions (5.2)–(5.6) on the applicability of our Hardy-Littlewood dissection, the condition in Lemma 5.3 will be satisfied in all cases of interest, so considering higher moments would provide no advantage within the current scope of our method.

We now describe the choice of parameters for an iterative process that concludes with a Hardy-Littlewood dissection. In view of (5.26), one has

$$(5.44) \qquad \int_{\mathfrak{m}_j} F_j(\boldsymbol{\alpha})|f_j(\boldsymbol{\alpha})|^{2s_j}\, d\boldsymbol{\alpha} \ll \widetilde{M}_j H_j \widetilde{H}_j P^{2-2\tau_j+\varepsilon} Q_j^{\lambda_{s_j}},$$

so when the minor arc contribution to the integral in (2.10) dominates, one computes the optimal values of $\phi_1, \ldots, \phi_j$ by first taking

$$\widetilde{M}_j \widetilde{H}_{j-1} P^2 Q_j^{\lambda_{s_j}} \approx \widetilde{M}_j H_j \widetilde{H}_j P^{2-2\tau_j} Q_j^{\lambda_{s_j}},$$

which reduces to $H_j^2 = P^{2\tau_j}$, or $k\phi_j = 1 - \tau_j$. Thus in (2.20) we take

$$\alpha_j = k^{-1}, \quad \beta_j = 0, \quad \text{and} \quad \gamma_j = 1 - \tau_j,$$

and we denote the resulting iterative process by D_j. Now suppose instead that the major arcs dominate. Then by Hölder's inequality one has

$$\int_{\mathfrak{M}_j} F_j(\boldsymbol{\alpha})|f_j(\boldsymbol{\alpha})|^{2s_j} d\boldsymbol{\alpha} \ll \mathcal{J}^{\frac{1}{k-j}} \left(\int_{\mathbb{T}^2} |f_j(\boldsymbol{\alpha})|^{2u_j} d\boldsymbol{\alpha} \right)^{\frac{k-j-1}{k-j}},$$

where $\mathcal{J}$ is the integral in the statement of Lemma 5.3, and where we have written

$$u_j = \frac{s_j(k-j) - 2}{k-j-1}.$$

It therefore follows from Lemma 5.3 that

$$(5.45) \qquad \int_{\mathfrak{M}_j} F_j(\boldsymbol{\alpha})|f_j(\boldsymbol{\alpha})|^{2s_j}\, d\boldsymbol{\alpha} \ll \widetilde{M}_j H_j \widetilde{H}_j P^{2+\varpi_j+\varepsilon} Q_j^{\Lambda_j},$$

where

$$\Lambda_j = \frac{4 - k - n + \lambda_{u_j}(k-j-1)}{k-j}.$$

A short computation shows that parameters are optimized in (2.10) when $H_j^2 P^{\varpi_j} \approx Q_j^{\delta_j}$, where $\delta_j = \lambda_{s_j} - \Lambda_j$. One therefore takes

$$\alpha_j = (\delta_j - 2k)^{-1}, \quad \beta_j = \delta_j, \quad \text{and} \quad \gamma_j = \delta_j - \varpi_j - 2$$

in (2.20), and in this case we denote the process by D_j'. One sees from (5.44) and (5.45) that the minor arcs dominate whenever

$$\delta_j(1 - \phi_1 - \cdots - \phi_j) \geq \varpi_j + 2\tau_j.$$

It should be emphasized that processes D_j and D_j' are only applicable when conditions (5.2)–(5.6) are satisfied. To overcome these rather serious restrictions, it may occasionally be useful to consider a one-dimensional Hardy-Littlewood dissection, in which the major and minor arcs are vertical strips defined only in terms of the rational approximations to α_k. This strategy has the advantage of being applicable in a large range of circumstances under which one of (5.2)–(5.6) fails to hold, but it turns out that the estimates produced are weaker than those stemming from the methods of §4 in all cases under consideration here.

6. Iterative Schemes and Permissible Exponents

Here we describe how the methods of the previous sections are combined to generate permissible exponents. Our major tools are the processes A_j, B_j, C_j, and D_j (and primed versions thereof) discussed in §§4-5. Occasionally, it transpires that the elementary process encoded in (2.22), wherein one takes $\phi_j = 1/K_j$, is the most effective applicable option, and we denote this process by E_j. Finally, it is often useful near the end of the iteration to observe that if the exponent λ_s is permissible and $2s \geq 3k$, then the exponent

$$\lambda_{s+1} = \max\{2s + 2 - k - n, \ \lambda_s + 2(1 - 2^{1-k})\}$$

is also permissible. This follows by dissecting into major and minor arcs and imitating the argument of the proof of Lemma 7.3 below, so we suppress the details here. We denote this process by D_0.

Our calculations often make use of the available permissible exponents for kth powers, which we denoted in §§4-5 by $\mu_s = \mu_s(k)$. We collect for reference in Table 6.1 the best such exponents currently available in the literature (see [42] for cubes, [11] for fourth powers, [32] for fifth and sixth powers, and [34] for seventh powers).

We further note that the exponent $\mu_2 = 2$ is always permissible for $k \geq 2$ and that the exponent $\mu_s = 2s - k$ is permissible for $s \geq s_0(k)$, where one has

$$s_0(3) = 4, \quad s_0(4) = 6, \quad s_0(5) = 9, \quad s_0(6) = 12, \quad \text{and} \quad s_0(7) = 17.$$

Our iterative processes for determining λ_s often depend on estimates for certain λ_u with $u > s$, and we must therefore establish preliminary permissible exponents before proceeding with such processes. For this purpose, we

TABLE 6.1. **Permissible exponents for kth powers**

k	s	μ_s	k	s	μ_s	k	s	μ_s
3	3	3.249414	6	4	4.333334	7	6	7.014382
4	3	3.183428	6	5	5.724697	7	7	8.541090
4	4	4.594193	6	6	7.231564	7	8	10.152633
4	5	6.213431	6	7	8.850572	7	9	11.846949
5	3	3.136258	6	8	10.560413	7	10	13.605568
5	4	4.438657	6	9	12.353671	7	11	15.424298
5	5	5.925080	6	10	14.203006	7	12	17.293221
5	6	7.541755	6	11	16.086042	7	13	19.192538
5	7	9.272729	7	3	3.063920	7	14	21.113930
5	8	11.077363	7	4	4.264118	7	15	23.052885
6	3	3.090910	7	5	5.589117	7	16	25.010539

note that the exponent

$$(6.1) \qquad \lambda_s = \min\{2s - 3, \mu_s(k)\}$$

is permissible for all $k > n \geq 1$ and $s \geq 3$, since one may either apply Theorem 3 of Wooley [37] or simply ignore the equation of degree n. The exponents $\lambda_2 = 2$ and $\lambda_3 = 3$ are easily seen to be best possible by considering diagonal solutions to the underlying system. For larger s, one may use (6.1) in conjunction with the process D_0 or E_j to obtain somewhat better preliminary bounds. After then estimating all moments of interest using the methods of §§2, 4, and 5, we may repeat the whole iteration, seeded with improved preliminary exponents. As this process is repeated, the resulting permissible values for λ_s form a bounded, decreasing (and hence convergent) sequence.

For example, a typical method for estimating λ_4 is to apply process A_1 with $K_1 = n$. If λ_4 is known to be permissible, then (4.5) shows that we can take $\phi_1 = (1 + \delta)/(3n + \delta)$, where $\delta = \lambda_4 - 4$. Since $\lambda_2 = 2$ is permissible, it then follows from Lemma 2.1 that the exponent $\lambda_4' = 4 + 2\phi_1$ is permissible, and we thus find that the converged exponent λ_4^* is the positive root of the equation

$$(6.2) \qquad \lambda^2 + (3n - 10)\lambda - (12n - 22) = 0.$$

A typical scheme for obtaining λ_5 and λ_6 would be to employ the processes

$$(6.3) \qquad F_0^2 f_0^6 \longrightarrow F_1 f_1^6 \implies (F_1^2)^{1/2}(f_1^{12})^{1/2}$$

and

$$(6.4) \qquad F_0^2 f_0^8 \longrightarrow F_1 f_1^8 \longrightarrow F_2 f_2^6 \implies (F_2^2)^{1/2}(f_2^{12})^{1/2}$$
$$\downarrow$$
$$f_1^{10}$$

TABLE 6.2. **Permissible exponents for $k = 4$ and $n = 1$**

s	Process	λ_s	ϕ_1	ϕ_2	ϕ_3
5	B_1'	6.000000	0.166667		
7	B_3': 01	9.369295	0.195614	0.1413	0.0022
8	D_0	11.119295			

TABLE 6.3. **Permissible exponents for $k = 4$ and $n = 2$**

s	Process	λ_s	ϕ_1	ϕ_2
5	B_2: 0	5.820313	0.164063	0.1250
6	B_2: 1	7.327240	0.164063	0.1250
7	B_2: 0	8.891122	0.173279	0.1250
8	B_2: 1	10.503055	0.176212	0.1250
9	B_2': 1	12.205039	0.184828	0.1376
10	D_0	14.000000		

in tandem. Whenever one obtains a new permissible λ_6 from (6.4), this value can be used in (6.3) to obtain an improved λ_5, and this in turn yields an improved λ_6 when substituted into (6.4). For intermediate moments, we typically use similar schemes involving the processes B_j and C_j, and for higher moments we are sometimes able to apply process D_j. In this latter case, the conditions (5.2)–(5.6) must be checked.

We now present our results for the values of k and n considered in Theorem 1.1. All values of λ_s and $\phi_1, \ldots, \phi_j$ given in the tables below are upper bounds, the converged values having been computed electronically using double-precision arithmetic. The processes listed are those used towards the end of the iteration. The sequence of numbers appearing after the process name in the second column is of the form $v_1 \cdots v_{j-1}(K_j)u$, where $v_i = t_i - s_i = s_i - s_{i+1}$ denotes the splitting parameter arising from the application of the Cauchy-Schwarz inequality in Lemma 2.2, and where $2u$ is the moment of F_j used in the process A_j'. Note that we always have $v_n = 0$ and that there are no parameters v_i to record when $j = 1$. When $K_j = k$, as is often the case, the (K_j) entry is omitted. When $K_j \neq k$, the value of ϕ_j should of course be compared with $1/K_j$ rather than $1/k$. To conserve space, we have omitted from the tables any values of s for which the best exponents λ_s are obtained simply by applying (6.1) or by using convexity (*i.e.*, Hölder's inequality) to interpolate between other available moments. We also often omit a series of rows in which a trivial process such as E_j or D_0 is applied repeatedly. It is a simple task for the reader to generate the missing rows if desired.

Our results for $k = 4$ are summarized in Tables 6.2–6.4. For the case $n = 1$, an application of process B_1' with $\phi_1 = 1/6$ yields the permissible

TABLE 6.4. **Permissible exponents for $k = 4$ and $n = 3$**

s	Process	λ_s	ϕ_1	ϕ_2	ϕ_3
4	A_1: (3)	4.274918	0.137459		
5	A_1': (3)5	5.605552	0.201851		
6	C_2: 1	7.027047	0.159179	0.1251	
7	C_2: 1	8.504521	0.166648	0.1251	
8	C_2: 1	10.042735	0.170048	0.1251	
9	C_2': 2	11.699788	0.184016	0.1303	
10	B_3: 11	13.445197	0.201583	0.1741	0.1251
11	D_0	15.195197			

TABLE 6.5. **Permissible exponents for $k = 5$ and $n = 1$**

s	Process	λ_s	ϕ_1	ϕ_2	ϕ_3	ϕ_4
6	A_3: 01	7.531036	0.144469	0.1249	0.0713	
7	B_3': 01	9.242834	0.163191	0.1551	0.1318	
8	B_4': 000	10.956428	0.168308	0.1529	0.1183	0.0409
9	B_4': 010	12.758315	0.173057	0.1536	0.1197	0.0439
10	B_4': 011	14.614641	0.183359	0.1653	0.1252	0.0561
11	D_0	16.489641				

TABLE 6.6. **Permissible exponents for $k = 5$ and $n = 2$**

s	Process	λ_s	ϕ_1	ϕ_2	ϕ_3	ϕ_4
5	B_2: 0	5.780000	0.130000	0.1000		
6	B_2: 1	7.249300	0.130000	0.1000		
7	B_2: 0	8.762642	0.136100	0.1000		
8	B_2: 1	10.318883	0.137999	0.1000		
9	B_2: 2	11.958978	0.145233	0.1000		
10	B_2: 3	13.753135	0.165216	0.1000		
11	B_4': 101	15.555313	0.176566	0.1618	0.1088	0.0079
12	B_4': 201(3)	17.418641	0.179666	0.1627	0.1103	0.0109
13	D_0	19.293641				

exponent $\lambda_5 = 6$, and an application of process B_3' is then sufficient to establish our claimed bound for $G^*(4, 1)$ in the next section. In the cases $n = 2$ and $n = 3$, we again rely heavily on the methods of §4. In the latter case, we see from (6.2) that $\lambda_4 = \frac{1}{2}(1 + \sqrt{57})$ is permissible, while in the former case one does better by simply averaging λ_3 and λ_5.

Tables 6.5–6.8 contain our results for $k = 5$. When $n = 1$, we obtain permissible exponents λ_4 and λ_5 simply by ignoring the underlying linear equation and referring to Table 6.1. We are then able to apply the methods of §4 when $6 \le s \le 10$, but the trivial dissection D_0 must be applied for $s \ge 11$. For the case $n = 2$, we use the process B_2 for $5 \le s \le 10$, the

TABLE 6.7. **Permissible exponents for $k = 5$ and $n = 3$**

s	Process	λ_s	ϕ_1	ϕ_2	ϕ_3
4	A_1: (3)	4.274918	0.137459		
5	A_1: (3)	5.605549	0.201850		
6	A_2: 1(3)	7.023173	0.130698	0.1886	
7	B_3: 11	8.501453	0.140107	0.1271	0.1000
8	B_3: 11	10.032891	0.144725	0.1318	0.1000
9	B_3: 11	11.619167	0.149058	0.1342	0.1000
10	B_3: 11	13.256636	0.153600	0.1368	0.1000
11	B_3: 11	14.940780	0.157695	0.1395	0.1000
12	B_3: 11	16.666958	0.161302	0.1417	0.1000
13	B_3': 11	18.440244	0.165518	0.1459	0.1056
14	B_3': 12	20.257290	0.170399	0.1512	0.1076
15	D_1	22.113201	0.175000		

TABLE 6.8. **Permissible exponents for $k = 5$ and $n = 4$**

s	Process	λ_s	ϕ_1	ϕ_2	ϕ_3	ϕ_4
4	A_1: (4)	4.196153	0.098077			
5	A_1: (4)	5.417173	0.139058			
6	A_2: 1(4)	6.776697	0.120850	0.1318		
7	A_2: 2(4)	8.165029	0.133957	0.1310		
8	C_3: 10	9.628974	0.136950	0.1280	0.1000	
9	C_3: 11	11.144940	0.143368	0.1290	0.1000	
10	C_3: 11	12.719540	0.147954	0.1339	0.1000	
11	C_3': 12	14.355945	0.154169	0.1403	0.1024	
12	B_4: 111	16.063314	0.162283	0.1520	0.1357	0.1000
13	B_4: 111	17.805813	0.167731	0.1565	0.1385	0.1000
14	B_4: 112	19.594823	0.171374	0.1618	0.1435	0.1000
15	D_1	21.414796	0.175000			
16	D_1	23.240729	0.175000			
17	D_1	25.092207	0.175000			

TABLE 6.9. **Permissible exponents for $k = 6$ and $n = 1$**

s	Process	λ_s	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5
8	A_4: 011	10.540066	0.133953	0.1274	0.1043	0.0486	
9	B_5': 0000	12.286025	0.141432	0.1354	0.1214	0.0886	0.0121
10	B_4': 013	14.068557	0.146129	0.1381	0.1196	0.0047	
11	B_5': 0110	15.908198	0.151408	0.1430	0.1264	0.0939	0.0245
12	B_5': 0111	17.783831	0.156912	0.1489	0.1307	0.1003	0.0344
13	B_5': 0112	19.698256	0.161386	0.1538	0.1384	0.1100	0.0464
14	B_5': 0123	21.626593	0.164296	0.1563	0.1394	0.1019	0.0059
15	D_0	23.564093					

TABLE 6.10. **Permissible exponents for $k = 6$ and $n = 2$**

s	Process	λ_s	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5
6	B_2: 1	7.175067	0.109792	0.0834			
7	B_2: 1	8.652662	0.112137	0.0834			
8	B_2': 1	10.175911	0.113411	0.0844			
9	B_2': 2	11.784059	0.121040	0.0878			
11	B_4': 102(5)	15.252389	0.143730	0.1366	0.1097	0.0224	
12	B_4': 202(5)	17.059819	0.147074	0.1385	0.1135	0.0325	
13	B_5': 0012	18.917400	0.154920	0.1599	0.1361	0.0915	0.0011
14	B_5': 1012	20.775271	0.156803	0.1602	0.1366	0.0924	0.0032
15	B_5': 2012(4)	22.671208	0.158249	0.1604	0.1370	0.0931	0.0047
16	E_2: 0	24.585627	0.161283	0.1667			
17	E_2: 1	26.505659	0.161929	0.1667			
18	E_2: 1	28.437055	0.162202	0.1667			
19	D_0	30.374555					

TABLE 6.11. **Permissible exponents for $k = 6$ and $n = 3$**

s	Process	λ_s	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5
4	A_1: (3)	4.274918	0.137459				
5	A_1: (3)	5.601860	0.200620				
6	A_2: 1(3)	7.007755	0.108971	0.1897			
7	B_3: 11	8.459328	0.115904	0.1059	0.0834		
8	B_3: 10	9.955773	0.118618	0.1091	0.0834		
9	B_3: 11	11.499301	0.121768	0.1105	0.0834		
10	B_3: 11	13.087094	0.125088	0.1126	0.0834		
11	B_3: 11	14.716154	0.128081	0.1145	0.0834		
12	B_3': 11	16.387686	0.131202	0.1172	0.0861		
13	B_3': 12	18.107582	0.135303	0.1222	0.0891		
14	B_3': 22	19.885634	0.141152	0.1229	0.0912		
15	B_3': 23	21.713856	0.147468	0.1319	0.0962		
16	B_5': 1102	23.586419	0.152999	0.1497	0.1461	0.0989	0.0141
17	B_5': 2201(3)	25.471260	0.155701	0.1478	0.1358	0.0864	0.0024
18	B_5': 2301(3)	27.404022	0.159250	0.1551	0.1397	0.0929	0.0137
19	E_3: 00	29.336863	0.161822	0.1634	0.1667		
20	D_0	31.274363					

process B_4' for $s = 11, 12$, and then the trivial dissection D_0 for $s \geq 13$. Note that the condition (5.2) precludes an application of the process D_j when $j \geq 1$. For the cases $n = 3$ and $n = 4$, we use process A_j in schemes resembling (6.3) and (6.4) to deal with small values of s. The methods B_j and C_j are applied in the intermediate ranges, and our first instance of a non-trivial Hardy-Littlewood dissection (process D_1) occurs towards the end. Note that by (6.2) we have $\lambda_4 = 3\sqrt{3} - 1$ in the case $n = 4$.

TABLE 6.12. **Permissible exponents for $k = 6$ and $n = 4$**

s	Process	λ_s	ϕ_1	ϕ_2	ϕ_3	ϕ_4
4	A_1: (4)	4.196153	0.098077			
5	A_1: (4)	5.418269	0.139423			
6	A_2: 1(4)	6.782118	0.100962	0.1333		
7	A_2: 2(4)	8.152910	0.111619	0.1326		
8	A_3: 12(4)	9.601812	0.113565	0.1094	0.1258	
9	B_4: 110	11.095219	0.120084	0.1127	0.1062	0.0834
10	B_4: 111	12.638762	0.123474	0.1178	0.1064	0.0834
11	B_4: 111	14.230087	0.127445	0.1209	0.1105	0.0834
12	B_4: 111	15.865735	0.131070	0.1240	0.1118	0.0834
13	B_4: 111	17.543932	0.134479	0.1271	0.1138	0.0834
14	B_4: 111	19.260946	0.137673	0.1298	0.1155	0.0834
15	B_4: 111	21.013655	0.140562	0.1323	0.1170	0.0834
16	B_4: 111	22.798679	0.143191	0.1345	0.1184	0.0834
17	B_4: 112	24.616144	0.145862	0.1372	0.1212	0.0834
18	B_4': 112	26.462938	0.148578	0.1401	0.1246	0.0894
19	B_4': 122	28.336199	0.151096	0.1430	0.1250	0.0904
20	B_4': 123	30.234207	0.153529	0.1461	0.1295	0.0923
22	D_1	34.072612	0.156251			
23	D_1	36.004439	0.156251			

TABLE 6.13. **Permissible exponents for $k = 6$ and $n = 5$**

s	Process	λ_s	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5
4	A_1: (5)	4.152068	0.076034				
5	A_1: (5)	5.310099	0.103367				
7	A_2: 2(5)	7.917810	0.106806	0.0992			
8	A_3: 12(5)	9.305039	0.108218	0.1050	0.0950		
9	C_4: 110(5)	10.732764	0.115071	0.1078	0.1029	0.1000	
10	C_4: 111(5)	12.215179	0.118278	0.1134	0.1029	0.1000	
11	C_4: 111	13.746822	0.122660	0.1165	0.1078	0.0834	
12	C_4: 111	15.326126	0.126463	0.1201	0.1091	0.0834	
13	C_4: 112	16.953368	0.130393	0.1239	0.1132	0.0834	
14	C_4': 122	18.628073	0.134584	0.1285	0.1139	0.0858	
15	B_5: 1111	20.347630	0.138780	0.1334	0.1260	0.1131	0.0834
16	B_5: 1111	22.102397	0.142146	0.1367	0.1288	0.1148	0.0834
17	B_5: 1111	23.891706	0.144952	0.1397	0.1315	0.1164	0.0834
18	B_5: 1112	25.712609	0.147759	0.1424	0.1343	0.1191	0.0834
19	B_5: 1122	27.561260	0.150298	0.1451	0.1367	0.1190	0.0834
20	D_2: 1	29.426174	0.151813	0.1459			
21	D_2: 1	31.305602	0.152495	0.1459			
22	D_2: 1	33.201039	0.153352	0.1459			
23	D_2: 1	35.109165	0.154225	0.1459			
24	D_2: 2	37.041143	0.155955	0.1459			

TABLE 6.14. **Permissible exponents for $k = 7$ and $n = 2$**

s	Process	λ_s	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6
7	B_2: 1	8.501481	0.096948	0.0715				
8	B_2: 2	10.037517	0.102461	0.0715				
9	A_4: 102	11.655991	0.109969	0.1038	0.0907	0.0212		
10	A_5: 1011	13.306492	0.115757	0.1113	0.1015	0.0778	0.0118	
11	A_5: 1012	15.004376	0.118864	0.1186	0.1108	0.0903	0.0290	
12	A_5: 1022	16.748268	0.123298	0.1244	0.1169	0.0932	0.0392	
13	B_5': 2002(4)	18.505071	0.125104	0.1208	0.1089	0.0826	0.0084	
14	B_5': 2012(4)	20.320411	0.128321	0.1266	0.1153	0.0919	0.0332	
15	B_6': 10110	22.168352	0.133117	0.1369	0.1263	0.1087	0.0745	0.0064
16	B_6': 10111	24.043725	0.135913	0.1423	0.1323	0.1128	0.0818	0.0190
17	E_2: 0	25.924955	0.136897	0.1429				
18	E_2: 1	27.819728	0.137077	0.1429				
28	E_2: 1	47.245560	0.141193	0.1429				
29	D_0	49.214310						

TABLE 6.15. **Permissible exponents for $k = 7$ and $n = 3$**

s	Process	λ_s	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6
6	A_2: 1(3)	6.989288	0.093742	0.1899				
7	B_3: 11	8.421812	0.099003	0.0911	0.0715			
8	B_3: 10	9.893690	0.100370	0.0929	0.0715			
9	B_3: 11	11.407212	0.102880	0.0939	0.0715			
10	B_3: 11	12.959305	0.105441	0.0956	0.0715			
11	B_3: 11	14.547900	0.107686	0.0971	0.0715			
12	B_3: 12	16.179415	0.110511	0.1003	0.0715			
13	B_3: 13	17.859992	0.114573	0.1060	0.0715			
14	B_3: 23	19.587208	0.119097	0.1060	0.0715			
15	B_3: 24	21.374285	0.124736	0.1149	0.0715			
16	B_6': 11000	23.204048	0.130256	0.1314	0.1392	0.1212	0.0891	0.0316
18	B_5': 2204(4)	26.884426	0.131322	0.1307	0.1273	0.1000	0.0035	
19	E_3: 11	28.775567	0.133635	0.1362	0.1429			
20	E_3: 00	30.671175	0.136232	0.1384	0.1429			
21	E_3: 10	32.578968	0.136369	0.1384	0.1429			
22	E_3: 11	34.499877	0.137200	0.1385	0.1429			
28	E_3: 11	46.206618	0.140557	0.1412	0.1429			
29	D_0	48.175368						

Our results for $k = 6$ are given in Tables 6.9–6.13. When $n = 1$, we use the exponents provided by Table 6.1 for $4 \leq s \leq 7$ before the processes of §4 apply. In the case $n = 2$, the method E_2 is applied for $16 \leq s \leq 18$, as the other processes suffer from the difficulties associated with the presence of an equation of small degree (see the discussion at the start of §3). Even for sextic-cubic systems, there is an application of E_3, and the trivial dissection

TABLE 6.16. **Permissible exponents for $k = 7$ and $n = 4$**

s	Process	λ_s	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6
4	A_1: (4)	4.196153	0.098077					
5	A_1: (4)	5.417964	0.139322					
7	A_2: 2(4)	8.143222	0.095655	0.1335				
8	A_3: 12(4)	9.576881	0.096879	0.0940	0.1277			
9	B_4: 110	11.045640	0.101891	0.0961	0.0908	0.0715		
10	B_4: 111	12.558885	0.104213	0.1001	0.0908	0.0715		
11	B_4: 110	14.113026	0.107229	0.1022	0.0941	0.0715		
12	B_4: 111	15.706561	0.109919	0.1047	0.0949	0.0715		
13	B_4: 111	17.338223	0.112538	0.1069	0.0966	0.0715		
14	B_4: 111	19.005452	0.115013	0.1090	0.0979	0.0715		
15	B_4: 111	20.706276	0.117311	0.1110	0.0991	0.0715		
16	B_4: 112	22.441402	0.119717	0.1135	0.1017	0.0715		
17	B_4: 122	24.209289	0.122259	0.1162	0.1017	0.0715		
18	B_4: 123	26.007751	0.124724	0.1191	0.1056	0.0715		
19	B_4: 124	27.837123	0.127267	0.1224	0.1108	0.0715		
20	B_4': 134	29.698634	0.130146	0.1265	0.1108	0.0717		
21	B_4': 234	31.588805	0.133078	0.1266	0.1111	0.0729		
22	B_6': 12301	33.502919	0.135647	0.1348	0.1315	0.1260	0.0849	0.0203
23	E_4: 111	35.430918	0.137357	0.1373	0.1384	0.1429		
24	E_4: 111	37.365975	0.138035	0.1382	0.1392	0.1429		
25	D_1	39.308782	0.138393					
26	D_1	41.252827	0.138393					
27	D_1	43.203549	0.138393					
28	D_1	45.155338	0.138393					
29	D_1	47.112879	0.138393					
30	D_1	49.071340	0.138393					
31	D_1	51.034758	0.138393					
32	D_1'	53.000000	0.138465					

D_0 must be applied for $s \geq 20$. Tables 6.12 and 6.13 illustrate a somewhat more robust set of processes working reasonably well for $n = 4$ and $n = 5$. In the former case, however, none of our methods perform any better than linear interpolation in the estimation of λ_{21}. At this stage, it would be natural to try to apply process D_2, but unfortunately the conditions (5.5)–(5.6) would mandate that $\phi_1 + \phi_2 \leq 1/4$, which is too restrictive to be useful.

Finally, we discuss the results for $k = 7$ presented in Tables 6.14–6.18. When $n \leq 3$, the usual difficulties associated with n being small relative to k appear again. In particular, our iterative methods allow us to conclude only that $G^*(7, 1) \leq 75$. Since this is inferior to the bound established in §3, we omit the corresponding table of exponents. When $n = 2$, we are forced to use the trivial method E_2 for $17 \leq s \leq 28$ and the process D_0 for $s \geq 29$.

TABLE 6.17. **Permissible exponents for $k = 7$ and $n = 5$**

s	Process	λ_s	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5
4	A_1: (5)	4.152068	0.076034				
5	A_1: (5)	5.310916	0.103639				
7	A_2: 2(5)	7.926446	0.092021	0.1000			
8	A_3: 12(5)	9.305738	0.093081	0.0907	0.0964		
9	C_4: 111(5)	10.719005	0.098168	0.0927	0.0891	0.1000	
10	A_4: 112(5)	12.181054	0.100678	0.0977	0.0918	0.1070	
11	B_5: 1111(5)	13.685200	0.104105	0.1003	0.0965	0.0882	0.1000
12	B_5: 1111	15.232318	0.107065	0.1035	0.0988	0.0918	0.0715
13	B_5: 1111	16.820282	0.110044	0.1063	0.1015	0.0928	0.0715
14	B_5: 1111	18.447836	0.112886	0.1091	0.1039	0.0947	0.0715
15	B_5: 1111	20.112796	0.115613	0.1117	0.1062	0.0960	0.0715
16	B_5: 1111	21.813081	0.118181	0.1142	0.1083	0.0974	0.0715
17	B_5: 1111	23.546275	0.120591	0.1165	0.1103	0.0986	0.0715
18	B_5: 1111	25.309919	0.122824	0.1187	0.1121	0.0997	0.0715
19	B_5: 1111	27.101472	0.124879	0.1207	0.1138	0.1007	0.0715
20	B_5: 1111	28.918411	0.126752	0.1225	0.1153	0.1016	0.0715
21	B_5: 1112	30.759389	0.128536	0.1243	0.1172	0.1034	0.0715
22	B_5: 1122	32.621829	0.130215	0.1260	0.1190	0.1034	0.0715
23	B_5: 1123	34.504025	0.131765	0.1278	0.1209	0.1060	0.0715
24	B_5: 1223	36.403928	0.133210	0.1294	0.1208	0.1059	0.0715
25	B_5: 1224	38.319537	0.134523	0.1310	0.1229	0.1090	0.0715
26	B_5: 1234	40.249299	0.135729	0.1324	0.1251	0.1089	0.0715
27	B_5': 1234	42.192454	0.136905	0.1340	0.1274	0.1133	0.0815
28	B_5': 1244	44.147468	0.138042	0.1355	0.1299	0.1136	0.0824
29	D_1	46.103320	0.138393				
30	D_1	48.064559	0.138393				
31	D_1	50.026521	0.138393				

For the $n = 3$ case, we must use process E_3 for $19 \leq s \leq 28$ and then D_0 for $s \geq 29$. In both of these cases, we include only the relative highlights in our tables. The remaining cases ($n = 4, 5, 6$) again illustrate the more robust nature of our methods when the sizes of k and n are somewhat comparable. As expected (compare with the tables in Vaughan and Wooley [32], [34]), mean value estimates are applied for small values of s, a Hardy-Littlewood dissection is admissible for larger s, and the number of differences peaks in the intermediate range. One could certainly express some disappointment, however, that the dissections D_j are not applicable in a wider range of circumstances. In the case $n = 5$, for example, one might hope to apply the process D_2 towards the end of the iteration, but the condition $\phi_1 + \phi_2 \leq 1/4$ arising from (5.6) renders this strategy ineffective, just as in the case $(k, n) = (6, 4)$. On the other hand, the process B_5 provides a reasonable

TABLE 6.18. **Permissible exponents for $k = 7$ and $n = 6$**

s	Process	λ_s	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6
5	A_1: (6)	5.242618	0.080873					
6	A_1': (6)7	6.495799	0.096551					
7	A_2: 2(6)	7.752343	0.088535	0.0785				
8	A_3: 12(6)	9.078034	0.089517	0.0875	0.0762			
9	A_4: 112(6)	10.436531	0.094402	0.0893	0.0863	0.0737		
10	A_4: 112(6)	11.845021	0.096818	0.0940	0.0886	0.0851		
11	A_4': 112(6)9	13.295588	0.100317	0.0965	0.0930	0.1006		
12	C_5: 1111(6)	14.790815	0.103310	0.0999	0.0954	0.0894	0.0834	
13	C_5: 1111	16.328402	0.106428	0.1027	0.0984	0.0907	0.0715	
14	C_5: 1111	17.907592	0.109390	0.1057	0.1009	0.0927	0.0715	
15	C_5: 1112	19.527069	0.112324	0.1086	0.1037	0.0954	0.0715	
16	C_5: 1122	21.185264	0.115185	0.1115	0.1067	0.0953	0.0715	
17	C_5: 1123(6)	22.880676	0.117983	0.1145	0.1100	0.0999	0.0834	
18	B_6: 11111	24.611260	0.120697	0.1174	0.1134	0.1076	0.0969	0.0715
19	B_6: 11111	26.372606	0.123104	0.1199	0.1158	0.1096	0.0982	0.0715
20	B_6: 11111	28.163038	0.125258	0.1222	0.1180	0.1115	0.0993	0.0715
21	B_6': 11111	29.979860	0.127284	0.1242	0.1200	0.1133	0.1004	0.0716
22	B_6: 11122	31.821013	0.129157	0.1262	0.1221	0.1153	0.1010	0.0715
23	B_6: 11222	33.683656	0.130859	0.1280	0.1239	0.1153	0.1009	0.0715
24	D_3: 11	35.564014	0.132257	0.1294	0.1251			
25	D_3: 11	37.457678	0.133222	0.1302	0.1251			
26	D_3: 11	39.364604	0.134013	0.1308	0.1251			
27	D_3: 11	41.283287	0.134808	0.1313	0.1251			
28	D_3: 11	43.212429	0.135517	0.1318	0.1251			
29	D_3: 11	45.150491	0.136127	0.1322	0.1251			
30	D_3: 12	47.099166	0.136844	0.1330	0.1251			
31	D_2: 1	49.057618	0.137704	0.1340				
32	D_1	51.022942	0.138393					

alternative to the Hardy-Littlewood dissection, both from a practical and a philosophical standpoint (see the remark at the end of §4).

In the next section, we indicate how to use the mean value estimates recorded in the above tables to establish the bounds quoted in Theorem 1.1.

7. THE CIRCLE METHOD

On recalling the definition of $G^*(k,n)$, we may suppose in establishing Theorem 1.1 that the system (1.1) possesses a non-singular real solution and a non-singular p-adic solution for every prime p and that for each i one has $c_i d_i \neq 0$. Moreover, after possibly replacing some of the variables x_i by $-x_i$ and then changing the signs of the corresponding coefficients in forms of odd degree, we may suppose that (1.1) has a non-singular real

solution $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_s)$ such that $\eta_i \geq 0$ for all i. It therefore suffices to consider the solubility of the system (1.1) in positive integers, so we define the exponential sums

$$F(\boldsymbol{\alpha}) = \sum_{1 \leq x \leq P} e(\alpha_k x^k + \alpha_n x^n) \quad \text{and} \quad f(\boldsymbol{\alpha}) = \sum_{x \in \mathcal{A}(P,R)} e(\alpha_k x^k + \alpha_n x^n).$$

Observe that $f(\boldsymbol{\alpha})$ is exactly as in (2.5), and that, in the notation of (2.2), $|F(\boldsymbol{\alpha})|^2 = F_0(\boldsymbol{\alpha})$. In this section and the next, we deviate from the notation of §2 by writing $F_i(\boldsymbol{\alpha}) = f(\mathbf{c}_i\boldsymbol{\alpha})$ and $f_i(\boldsymbol{\alpha}) = f(\mathbf{c}_i\boldsymbol{\alpha})$, where $\mathbf{c}_i\boldsymbol{\alpha}$ denotes the vector $(c_i\alpha_k, d_i\alpha_n)$. Now let t be a parameter at our disposal, write $s = 2u + t$, and define

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

Then one sees by orthogonality that

(7.1) $$N(P) = \int_{\mathbb{T}^2} \mathcal{F}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha}$$

is the number of solutions of the system (1.1) with the variables satisfying

$$1 \leq x_i \leq P \quad (i = 1, \ldots, t) \quad \text{and} \quad x_i \in \mathcal{A}(P, R) \quad (i = t+1, \ldots, s).$$

We now describe our Hardy-Littlewood dissection. Write

$$C_k = \max_{1 \leq i \leq s} |c_i|, \quad C_n = \max_{1 \leq i \leq s} |d_i|, \quad \text{and} \quad X_i = 2k^2 C_i P^{i-1} \; (i = k, n).$$

When $n > 1$, we define the major arcs $\mathfrak{M}$ to be the union of the rectangles

(7.2) $$\mathfrak{M}(q, \mathbf{a}) = \{\boldsymbol{\alpha} \in \mathbb{T}^2 : |q\alpha_i - a_i| < X_i^{-1} \; (i = k, n)\}$$

with $0 \leq a_k, a_n \leq q \leq P$ and $(q, a_k, a_n) = 1$, and write $\mathfrak{m} = \mathbb{T}^2 \setminus \mathfrak{M}$ for the minor arcs. When $n = 1$, we use the same definition, except that $\mathfrak{M}$ is further restricted to those $\mathfrak{M}(q, \mathbf{a})$ for which $(q, a_k) \leq P^\varepsilon$. Note that the $\mathfrak{M}(q, \mathbf{a})$ may not be disjoint when $n = 1$.

We say that Δ_s is an admissible exponent if the exponent $\lambda_s = 2s - (k + n) + \Delta_s$ is permissible. The following lemma tells us that the minor arc contribution to the integral (7.1) is small whenever t is sufficiently large in terms of Δ_u.

Lemma 7.1. *Suppose that Δ_u is an admissible exponent and that $t > 2^{k-1}\Delta_u$, and further write $s = 2u + t$. Then for some $\delta > 0$, one has*

$$\int_{\mathfrak{m}} \mathcal{F}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \ll P^{s-k-n-\delta}.$$

Proof. By applying Hölder's inequality and considering the underlying diophantine equations, one finds that

$$\int_{\mathfrak{m}} \mathcal{F}(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} \ll \sup_{\boldsymbol{\alpha}\in\mathfrak{m}} |F_i(\boldsymbol{\alpha})|^t \int_{\mathbb{T}^2} |f(\boldsymbol{\alpha})|^{2u}\,d\boldsymbol{\alpha}$$

for some i with $1 \le i \le t$. Now by applying Baker [2], Theorem 5.1, and arguing as in the proof of Wooley [37], Lemma 7.4, we see that

$$\sup_{\boldsymbol{\alpha}\in\mathfrak{m}} |F_i(\boldsymbol{\alpha})| \ll P^{1-2^{1-k}+\varepsilon},$$

and the lemma follows immediately. $\square$

We choose u and t according to the Table 7.1, so as to minimize the value of $s = 2u + t$ for which Lemma 7.1 applies. Note that bounds for $G^*(k,n)$ in cases appearing in Theorem 1.1 but omitted from the table have been established in §3 or have been quoted from other sources.

TABLE 7.1. **Choices for (u,t) in the decomposition $s = 2u + t$**

	1	2	3	4	5	6
4	$(7,3)$	$(9,2)$	$(9,6)$			
5	$(10,10)$	$(11,9)$	$(15,2)$	$(16,4)$		
6	$(13,23)$	$(18,14)$	$(19,11)$	$(22,3)$	$(23,4)$	
7		$(27,18)$	$(27,16)$	$(30,5)$	$(31,2)$	$(31,4)$

One can verify from the tables in §6 that the hypothesis of Lemma 7.1 is satisfied for each pair (u,t) in Table 7.1. It now suffices to show that

$$(7.3) \qquad \int_{\mathfrak{M}} \mathcal{F}(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} \gg P^{s-k-n},$$

and this follows by a relatively straightforward extension of the argument of Wooley [37]. The basic strategy is to use the t variables ranging over a complete interval to prune back to a thinner set of major arcs, on which we can obtain asymptotics for $f_i(\boldsymbol{\alpha})$ that allow us to make use of local information. Let us introduce the notation

$$S_i(q,\mathbf{a}) = \sum_{x=1}^{q} e((c_i a_k x^k + d_i a_n x^n)/q),$$

$$v_i(\boldsymbol{\beta}) = \int_0^P e(c_i \beta_k \gamma^k + d_i \beta_n \gamma^n)\,d\gamma,$$

and

$$w_i(\boldsymbol{\beta}) = \int_R^P \rho\left(\frac{\log\gamma}{\log R}\right) e(c_i \beta_k \gamma^k + d_i \beta_n \gamma^n)\,d\gamma,$$

where ρ denotes Dickman's function (see for example Vaughan [30], §12.1). We first recall some standard estimates for these functions. It follows easily from Theorem 7.1 of Vaughan [30] that

$$(7.4) \qquad S_i(q, \mathbf{a}) \ll (q, c_i a_k, d_i a_n)^{1/k} q^{1-1/k+\varepsilon}.$$

Moreover, by applying the argument of Vaughan [30], Theorem 7.3, as in the proof of Wooley [37], Lemma 8.6, one finds that

$$(7.5) \qquad v_i(\boldsymbol{\beta}) \ll P(1 + P^n|\beta_n| + P^k|\beta_k|)^{-1/k}$$

and

$$(7.6) \qquad w_i(\boldsymbol{\beta}) \ll P(1 + P^n|\beta_n| + P^k|\beta_k|)^{-1/k},$$

since we have assumed that $c_i d_i \neq 0$ for each i. Now let $W \leq R$ be a parameter at our disposal. We define the pruned major arcs $\mathfrak{N}$ to be the union of the rectangles

$$(7.7) \qquad \mathfrak{N}(q, \mathbf{a}) = \{\boldsymbol{\alpha} \in \mathbb{T}^2 : |\alpha_i - a_i/q| < WP^{-i} \ \ (i = k, n)\}$$

with $0 \leq a_k, a_n \leq q \leq W$ and $(q, a_k, a_n) = 1$. Note here that the condition $(q, a_k) \leq P^\varepsilon$ is automatically satisfied, since $q \leq R$. Furthermore, the $\mathfrak{N}(q, \mathbf{a})$ are pairwise disjoint, even when $n = 1$. We need the following easy extension of Wooley [37], Lemma 9.2, in order to carry out our pruning argument.

Lemma 7.2. *If $c_i d_i \neq 0$ and T is a real number with $T > 3k$, then one has*

$$\int_{\mathfrak{M}} |F_i(\boldsymbol{\alpha})|^T \, d\boldsymbol{\alpha} \ll P^{T-k-n}$$

and, for some $\sigma > 0$,

$$\int_{\mathfrak{M}\backslash\mathfrak{N}} |F_i(\boldsymbol{\alpha})|^T \, d\boldsymbol{\alpha} \ll P^{T-k-n} W^{-\sigma}.$$

Proof. Suppose that $T > 3k$. When $\boldsymbol{\alpha} \in \mathfrak{M}(q, \mathbf{a})$, we write $\beta_i = \alpha_i - a_i/q$ for $i = k, n$ and define the function $V_i(\boldsymbol{\alpha}) = q^{-1} S_i(q, \mathbf{a}) v_i(\boldsymbol{\beta})$. In order to make this well-defined when $n = 1$, we can associate $\boldsymbol{\alpha}$ to the $\mathfrak{M}(q, \mathbf{a})$ having minimal q, since the arcs corresponding to a fixed q are pairwise disjoint. Then when $\boldsymbol{\alpha} \in \mathfrak{M}(q, \mathbf{a}) \subseteq \mathfrak{M}$, we have by Lemma 4.4 of Baker [2] that

$$|F_i(\boldsymbol{\alpha})|^T \ll |V_i(\boldsymbol{\alpha})|^T + P^\varepsilon (q^{1-1/k})^T.$$

Write $\mathfrak{W}$ for either $\mathfrak{M}$ or $\mathfrak{M}\backslash\mathfrak{N}$. Then one sees easily from (7.2) that

$$\int_{\mathfrak{W}} P^\varepsilon (q^{1-1/k})^T \, d\boldsymbol{\alpha} \ll P^{2-k-n+\varepsilon} \sum_{q \leq P} q^{T(1-1/k)} \ll P^{T-k-n-\delta}$$

for some $\delta > 0$, since $T > 3k$. Now by (7.4), one has

$$(7.8) \qquad \int_{\mathfrak{W}} |V_i(\boldsymbol{\alpha})|^T \, d\boldsymbol{\alpha} \ll \sum_{q \leq P} \sum_{\mathbf{a} \in [0,q]^2} q^{-T/k + \varepsilon} \int_{\mathfrak{W}(q,\mathbf{a})} |v_i(\boldsymbol{\beta})|^T \, d\boldsymbol{\alpha},$$

where have written $\mathfrak{W}(q, \mathbf{a}) = \mathfrak{M}(q, \mathbf{a})$ when $\mathfrak{W} = \mathfrak{M}$ and $\mathfrak{W}(q, \mathbf{a}) = \mathfrak{M}(q, \mathbf{a}) \setminus \mathfrak{N}(q, \mathbf{a})$ when $\mathfrak{W} = \mathfrak{M} \setminus \mathfrak{N}$. Now set $Y = 1$ if $\mathfrak{W} = \mathfrak{M}$ or if $q > W$, and put $Y = W$ otherwise. Then by applying (7.5) and making a change of variable, one finds that

$$\int_{\mathfrak{W}(q,\mathbf{a})} |v_i(\boldsymbol{\beta})|^T \, d\boldsymbol{\alpha} \ll P^T \int_{\mathfrak{W}(q,\mathbf{a})} (1 + P^n |\beta_n|)^{-T/2k} (1 + P^k |\beta_k|)^{-T/2k} \, d\boldsymbol{\alpha}$$
$$\ll P^{T-k-n} Y^{1-T/2k}.$$

Thus on writing $Z = 1$ if $\mathfrak{W} = \mathfrak{M}$ and $Z = W$ if $\mathfrak{W} = \mathfrak{M} \setminus \mathfrak{N}$, we obtain from (7.8) that

$$\int_{\mathfrak{W}} |V_i(\boldsymbol{\alpha})|^T \, d\boldsymbol{\alpha} \ll P^{T-k-n} \left(\sum_{q \leq W} Z^{1-T/2k} q^{2-T/k+\varepsilon} + \sum_{q > W} q^{2-T/k+\varepsilon} \right)$$
$$\ll P^{T-k-n} \left(Z^{1-T/2k} + W^{3-T/k+\varepsilon} \right),$$

and the lemma follows. $\square$

When t is as in Table 7.1, write $r = t$ if t is even and $r = t + 1$ if t is odd. The argument given in the following lemma underlies the process D_0 discussed in §6 and will be important in our pruning argument.

Lemma 7.3. *If* $t > 2^{k-1} \Delta_u$ *and* $2u + r > 3k$, *then one has*

$$\int_{\mathbb{T}^2} |f(\boldsymbol{\alpha})|^{2u+r} \, d\boldsymbol{\alpha} \ll P^{2u+r-k-n}.$$

Proof. Write $\mathcal{I}$ for the integral in question. Since r is even, one sees by considering the underlying diophantine equations that

$$\mathcal{I} \leq \int_{\mathbb{T}^2} |F(\boldsymbol{\alpha})|^r |f(\boldsymbol{\alpha})|^{2u} \, d\boldsymbol{\alpha},$$

and we now dissect into major and minor arcs as in (7.2). Since $r \geq t$, we have as in the proof of Lemma 7.1 that

$$\int_{\mathfrak{m}} |F(\boldsymbol{\alpha})|^r |f(\boldsymbol{\alpha})|^{2u} \, d\boldsymbol{\alpha} \ll P^{2u+r-k-n-\delta}$$

for some $\delta > 0$. We may therefore suppose that the contribution from the major arcs dominates, in which case we have by Hölder's inequality that

$$\mathcal{I} \ll \int_{\mathfrak{M}} |F(\boldsymbol{\alpha})|^r |f(\boldsymbol{\alpha})|^{2u} \, d\boldsymbol{\alpha} \ll \left(\int_{\mathfrak{M}} |F(\boldsymbol{\alpha})|^{2u+r} \, d\boldsymbol{\alpha} \right)^{\frac{r}{2u+r}} \mathcal{I}^{\frac{2u}{2u+r}},$$

and the result now follows from Lemma 7.2. $\qquad\square$

We are now able to achieve our pruning of the major arcs. By applying the trivial inequality $|z_1 \cdots z_m| \le |z_1|^m + \cdots + |z_m|^m$, we find that

$$\int_{\mathfrak{M}\backslash\mathfrak{N}} \mathcal{F}(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} \ll \int_{\mathfrak{M}\backslash\mathfrak{N}} |F_i(\boldsymbol{\alpha})^t f_j(\boldsymbol{\alpha})^{2u}|\,d\boldsymbol{\alpha}$$

for some i and j with $1 \le i \le t$ and $t+1 \le j \le s$. Then by applying Hölder's inequality and considering the underlying diophantine equations, we obtain

$$\int_{\mathfrak{M}\backslash\mathfrak{N}} \mathcal{F}(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} \ll \left(\int_{\mathfrak{M}\backslash\mathfrak{N}} |F_i(\boldsymbol{\alpha})|^T\,d\boldsymbol{\alpha}\right)^{\frac{r}{2u+r}} \left(\int_{\mathbb{T}^2} |f(\boldsymbol{\alpha})|^{2u+r}\,d\boldsymbol{\alpha}\right)^{\frac{2u}{2u+r}},$$

where $T = (2u+r)tr^{-1}$, and where r is as in the paragraph preceding Lemma 7.3. In all cases under consideration, we see from Table 7.1 that $t \ge 2$ and $2u + r > 4k$, and hence $T \ge \frac{3}{4}(2u + r) > 3k$. On recalling that $s = 2u + t$, we therefore deduce from Lemmas 7.2 and 7.3 that

$$(7.9) \qquad \int_{\mathfrak{M}\backslash\mathfrak{N}} \mathcal{F}(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} \ll P^{s-k-n}W^{-\sigma}$$

for some $\sigma > 0$. To establish (7.3), it now suffices to show that

$$(7.10) \qquad \int_{\mathfrak{N}} \mathcal{F}(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} \gg P^{s-k-n},$$

and for this we may follow the argument of Wooley [37], §10, fairly closely. We obtain some simplifications, however, as a result of our definition (7.7), in which the dimensions of the boxes $\mathfrak{N}(q, \mathbf{a})$ are independent of q.

When $\boldsymbol{\alpha} \in \mathfrak{N}(q, \mathbf{a})$, write $\beta_i = \alpha_i - a_i/q$ for $i = k, n$, define $V_i(\boldsymbol{\alpha})$ as in the proof of Lemma 7.2, and write $W_i(\boldsymbol{\alpha}) = q^{-1}S_i(q, \mathbf{a})w_i(\boldsymbol{\beta})$. Then, whenever $\boldsymbol{\alpha} \in \mathfrak{N}(q, \mathbf{a}) \subseteq \mathfrak{N}$, we see from Lemma 8.5 of Wooley [37] that

$$f_i(\boldsymbol{\alpha}) - W_i(\boldsymbol{\alpha}) \ll W^2P(\log P)^{-1}$$

and from Theorem 7.2 of Vaughan [30] that

$$F_i(\boldsymbol{\alpha}) - V_i(\boldsymbol{\alpha}) \ll W^2.$$

In view of (7.7), we have $\mathrm{meas}(\mathfrak{N}) \ll W^5 P^{-k-n}$, and we therefore find as in the argument of [37], Lemma 10.1, that

$$\int_{\mathfrak{N}} \mathcal{F}(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} = \int_{\mathfrak{N}} \prod_{i=1}^{t} V_i(\boldsymbol{\alpha}) \prod_{i=t+1}^{s} W_i(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} + O(P^{s-k-n}W^7(\log P)^{-1}).$$

As usual, we can factor the integral on the right as a product of a truncated singular series and a truncated singular integral. On writing

$$(7.11) \qquad S(q) = \sum_{\substack{a=1 \\ (q,a,b)=1}}^{q} \sum_{b=1}^{q} \prod_{i=1}^{s} q^{-1} S_i(q, a, b), \qquad \mathfrak{S}(W) = \sum_{q \le W} S(q),$$

and

$$J(W) = \int_{-WP^{-k}}^{WP^{-k}} \int_{-WP^{-n}}^{WP^{-n}} \prod_{i=1}^{t} v_i(\boldsymbol{\beta}) \prod_{i=t+1}^{s} w_i(\boldsymbol{\beta}) \, d\boldsymbol{\beta},$$

we see immediately that

$$(7.12) \qquad \int_{\mathfrak{N}} \mathcal{F}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} = \mathfrak{S}(W)J(W) + O(P^{s-k-n}W^7(\log P)^{-1}).$$

We now complete $\mathfrak{S}(W)$ and $J(W)$ to infinity by defining

$$\mathfrak{S} = \sum_{q=1}^{\infty} S(q) \qquad \text{and} \qquad J = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \prod_{i=1}^{t} v_i(\boldsymbol{\beta}) \prod_{i=t+1}^{s} w_i(\boldsymbol{\beta}) \, d\boldsymbol{\beta}.$$

By applying the argument of Vaughan [30], Lemma 2.11, as in Lemmas 10.4 and 10.5 of Wooley [37], we find that the function $S(q)$ defined in (7.11) is multiplicative. Moreover, since $c_i d_i \ne 0$ for each i, we have $(q, c_i a, d_i b) \ll (q, a, b)$, and it then follows from (7.4) that

$$(7.13) \qquad\qquad S(q) \ll q^{2-s/k+\varepsilon}.$$

When $s > 3k$, we therefore have that $\mathfrak{S}$ is absolutely convergent and that

$$\mathfrak{S} - \mathfrak{S}(W) \ll W^{-\tau}$$

for some $\tau > 0$. Furthermore, on using (7.5) and (7.6) and making a change of variable, we find that

$$(7.14) \qquad J \ll P^{s-k-n} \int_0^{\infty} \int_0^{\infty} (1 + \beta_k)^{-s/2k}(1 + \beta_n)^{-s/2k} d\boldsymbol{\beta} \ll P^{s-k-n}$$

whenever $s > 2k$. Thus J is absolutely convergent, and it follows similarly that

$$J - J(W) \ll P^{s-k-n}W^{1-s/2k},$$

again provided that $s > 2k$. On taking $W = (\log P)^{1/8}$, we therefore deduce from (7.12) that

$$(7.15) \qquad\qquad \int_{\mathfrak{N}} \mathcal{F}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} = \mathfrak{S}J + O(P^{s-k-n}W^{-\sigma})$$

for some $\sigma > 0$. The following lemma provides us with the required asymptotic information concerning the singular integral J.

Lemma 7.4. *If $s > 2k$, then one has $J \gg P^{s-k-n}$.*

Proof. Here we deviate from the argument of [37] and instead apply a method of Schmidt [26], which avoids the use of Fourier's integral formula. Let T be a positive real number, and introduce the functions

$$K_T(\beta) = \left(\frac{\sin \pi \beta T^{-1}}{\pi \beta T^{-1}} \right)^2 \quad \text{and} \quad \mathcal{K}_T(\boldsymbol{\beta}) = K_T(\beta_k) K_T(\beta_n).$$

It follows from Lemma 14.1 of Baker [2] that

$$(7.16) \qquad \widehat{K}_T(y) = \int_{-\infty}^{\infty} K_T(\beta) \, e(\beta y) \, d\beta = T \max(0, 1 - T|y|)$$

for all real numbers y. We write

$$J_T = \int_{\mathbb{R}^2} \mathcal{K}_T(\boldsymbol{\beta}) \prod_{i=1}^{t} v_i(\boldsymbol{\beta}) \prod_{i=t+1}^{s} w_i(\boldsymbol{\beta}) \, d\boldsymbol{\beta}$$

and note that the argument leading to (7.14) gives

$$(7.17) \quad J - J_T \ll P^s \int_{\mathbb{R}^2} (1 - \mathcal{K}_T(\boldsymbol{\beta}))(1 + P^k|\beta_k|)^{-\frac{s}{2k}}(1 + P^n|\beta_n|)^{-\frac{s}{2k}} \, d\boldsymbol{\beta}.$$

A simple calculation reveals that

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

so on making a change of variable in (7.17) and considering the resulting integral over the regions $|\boldsymbol{\beta}| \le T$ and $|\boldsymbol{\beta}| > T$ separately, it is easily shown that

$$J - J_T \ll P^{s-k-n} T^{-\delta}$$

for some $\delta > 0$, provided that $s > 2k$. Hence for any fixed P, we have

$$(7.18) \qquad\qquad\qquad J = \lim_{T \to \infty} J_T,$$

and so it suffices to analyze J_T. By making a change of variable, we find that

$$(7.19) \qquad J_T = P^s \int_{\mathfrak{B}} \mathcal{H}(\boldsymbol{\gamma}) \, \widehat{K}_T(P^k F(\boldsymbol{\gamma})) \, \widehat{K}_T(P^n G(\boldsymbol{\gamma})) \, d\boldsymbol{\gamma},$$

where we have written

$$F(\boldsymbol{\gamma}) = c_1 \gamma_1^k + \cdots + c_s \gamma_s^k, \qquad G(\boldsymbol{\gamma}) = d_1 \gamma_1^n + \cdots + d_s \gamma_s^n,$$

$$\mathcal{H}(\boldsymbol{\gamma}) = \prod_{i=t+1}^{s} \rho\left(\frac{\log(P \gamma_i)}{\log R} \right), \qquad \text{and} \qquad \mathfrak{B} = [0,1]^t \times [R/P, 1]^{2u}.$$

Since we have assumed that the system $F(\boldsymbol{\gamma}) = G(\boldsymbol{\gamma}) = 0$ possesses a non-singular real solution $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_s)$ with $\eta_i \ge 0$ for each i, the Implicit Function Theorem ensures that locally near $\boldsymbol{\eta}$ there is an $(s-2)$-dimensional

space of real solutions, continuously parameterized by $s - 2$ of the coordinates. Therefore, by exploiting continuity as in the proof of [37], Lemma 6.2, we may suppose that each η_i is non-zero and hence that $\boldsymbol{\eta}$ lies in the interior of $\mathfrak{B}$ for P sufficiently large. Now consider the map $\varphi : \mathbb{R}^s \to \mathbb{R}^s$ defined by

$$\varphi_1(\boldsymbol{\gamma}) = F(\boldsymbol{\gamma}), \quad \varphi_2(\boldsymbol{\gamma}) = G(\boldsymbol{\gamma}), \quad \text{and} \quad \varphi_i(\boldsymbol{\gamma}) = \gamma_i \quad (3 \leq i \leq s).$$

By the Inverse Function Theorem, there is an open set $U \subseteq \mathfrak{B}$ containing $\boldsymbol{\eta}$, and an open set V containing $(0, 0, \eta_3, \ldots, \eta_s)$, such that φ maps U injectively onto V. Since $\mathcal{H}(\boldsymbol{\gamma}) \gg 1$ on $\mathfrak{B}$ and the integrand in (7.19) is non-negative, we have by a change of variable that

$$(7.20) \qquad J_T \gg P^{s-k-n} \int_{V^*} \widehat{K}_T(u_1) \widehat{K}_T(u_2) \, du_1 \cdots du_s,$$

where V^* is the set of all $\mathbf{u}$ for which $(P^{-k}u_1, P^{-n}u_2, u_3, \ldots, u_s) \in V$. In particular, it is clear that the projection of V^* onto the first two components contains the set $\mathfrak{D} = [-1/2T, 1/2T]^2$ whenever $T \geq 1$ and P is sufficiently large. By (7.16), the integrand in (7.20) is bounded below on $\mathfrak{D}$ by $(T/2)^2$, and $\mathrm{meas}(\mathfrak{D}) \gg T^{-2}$, so it follows immediately that $J_T \gg P^{s-k-n}$ for $T \geq 1$, where the implicit constant is independent of T. The lemma therefore follows from (7.18) on letting $T \to \infty$. $\qquad \square$

We now turn our attention to the singular series.

Lemma 7.5. *If $s > 3k$, then one has $\mathfrak{S} > 0$.*

Proof. On recalling (7.13), we see that the series

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

is absolutely convergent whenever $s > 3k$. We therefore find as in Wooley [37], Lemma 10.8, that $\mathfrak{S}$ is represented by the absolutely convergent product

$$\mathfrak{S} = \prod_p T(p),$$

and that there exists an integer p_0 such that

$$\frac{1}{2} \leq \prod_{p \geq p_0} T(p) \leq \frac{3}{2}.$$

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

$$c_1 x_1^k + \cdots + c_s x_s^k \equiv d_1 x_1^n + \cdots + d_s x_s^n \equiv 0 \pmod{q}.$$

By applying the argument of [30], Lemma 2.12, as in [37], Lemma 10.9, we find that

$$\sum_{d|q} S(d) = q^{2-s} M_s(q),$$

and it follows that

$$T(p) = \lim_{h\to\infty} \sum_{d|p^h} S(d) = \lim_{h\to\infty} p^{h(2-s)} M_s(p^h).$$

Since we have assumed that the system (1.1) possesses a non-singular p-adic solution for each prime p, we may apply a Hensel's Lemma argument as in Wooley [37], Lemma 6.7, to conclude that there exists an integer $u = u(p) < \infty$ such that for all h one has

$$M_s(p^h) \geq p^{(h-u)(s-2)}.$$

It follows that $T(p) \geq p^{u(2-s)}$ for each $p < p_0$, and this completes the proof. $\qquad\Box$

In view of Lemmas 7.4 and 7.5, the lower bound (7.3) follows immediately from (7.9) and (7.15). Thus on recalling Lemma 7.1, we find that

$$(7.21) \qquad G^*(k,n) \leq \min_{u\in\mathbb{N}} \left(2u + \left[2^{k-1}\Delta_u\right] + 1\right),$$

where Δ_u denotes an admissible exponent. The bounds claimed in Theorem 1.1 now follow by choosing u as indicated in Table 7.1 and examining the results of §6.

8. Dealing with Zero Coefficients

Here we prove Theorem 1.2 as an indication of how one may adapt the argument of §7 to handle systems in which a limited number of zero coefficients are present. Although Theorem 1.2 is stated for systems consisting of a quintic form and a cubic form,

$$(8.1) \qquad c_1 x_1^5 + \cdots + c_s x_s^5 = d_1 x_1^3 + \cdots + d_s x_s^3 = 0,$$

it is certainly possible to establish analogous results for other pairs of degrees. Indeed, our argument closely follows that of Wooley [37] for cubic-quadratic systems. It is somewhat awkward, however, to write down a general statement in terms of k and n, since the choice of parameters appearing in conditions (a), (b), and (d) depends on the available results concerning the solubility of single additive equations of degree k and n.

One knows from Baker [3] that an additive cubic equation in 7 or more variables has a non-trivial integral solution. The following result on quintics is well-known to experts but not readily found in the literature.

Lemma 8.1. *Suppose that $s \geq 17$ and that $c_1, \ldots, c_s$ are integers. Then the equation*

$$c_1 x_1^5 + \cdots + c_s x_s^5 = 0$$

has a non-trivial integral solution.

Proof. The existence of non-trivial p-adic solutions for each prime p is ensured by Gray [17] whenever $s \geq 16$. The result therefore follows by applying the Hardy-Littlewood method through an adaptation of the arguments of Vaughan and Wooley [31], [32]. $\square$

Cases (ii) and (iii) of Theorem 1.2 can now be dealt with quite easily. The following observation, modeled after [37], Lemma 6.3, suffices.

Lemma 8.2. *Suppose that the system* (8.1) *satisfies conditions* (a), (b), *and* (c) *of Theorem* 1.2. *Further suppose that at least* 16 *of the d_i are zero or that at least* 6 *of the c_i are zero. Then the equations* (8.1) *have a non-trivial integral solution.*

Proof. Suppose that at least 16 of the d_i are zero. Then by a rearrangement of the variables, we may suppose that $d_1 = \cdots = d_{16} = 0$. It then follows from condition (a) of the theorem that $s \geq 23$, and thus we know from Baker [3] that there are integers $y_{17}, \ldots, y_s$, not all zero, such that

$$d_{17} y_{17}^3 + \cdots + d_s y_s^3 = 0.$$

Now write $C_i = c_i$ for $i = 1, \ldots, 16$ and $C_{17} = c_{17} y_{17}^5 + \cdots + c_s y_s^5$. Then by Lemma 8.1, there are integers $u_1, \ldots, u_{17}$, not all zero, satisfying the equation

$$C_1 u_1^5 + \cdots + C_{17} u_{17}^5 = 0.$$

The s-tuple $\mathbf{x} = (u_1, \ldots, u_{16}, y_{17} u_{17}, \ldots, y_s u_{17})$ is now a non-trivial integral solution to the system (8.1). A similar argument, invoking condition (b) of the theorem, applies to the case where at least 6 of the c_i are zero. $\square$

Theorem 1.2 now follows in cases (ii) and (iii). Note that the infinitude of solutions asserted by the theorem follows immediately from the existence of a single non-trivial solution, in view of the homogeneity of the system.

Notice that condition (c) of the theorem was not actually needed to establish Lemma 8.2. In order to obtain a suitable analogue of the lemma when k or n is even, however, one not only needs condition (c) but also some additional maneuvers (see for example [37], Lemma 6.4) to ensure the solubility of the auxiliary equations occurring in the proof.

In order to establish Theorem 1.2, we may suppose from now on that $s \geq 32$, that at most 15 of the coefficients d_i are zero, and that at most 5 of the coefficients c_i are zero. By rearranging variables, we may further

suppose that the first m of the d_i are zero and that the last n of the c_i are zero. On writing $h = s - m - n$, we then have

(8.2) $$m \le 15, \qquad n \le 5, \qquad \text{and} \qquad h \ge 12.$$

Let $f_i(\boldsymbol{\alpha})$ and $F_i(\boldsymbol{\alpha})$ be as in §7 (with $k = 5$ and $n = 3$), and write

$$\mathcal{F}(\boldsymbol{\alpha}) = \prod_{i=1}^{m+h-2} f_i(\boldsymbol{\alpha}) \prod_{i=m+h-1}^{s} F_i(\boldsymbol{\alpha})$$

We also introduce the shorthand notation

$$g_i(\alpha_5) = f_i(\alpha_5, 0) \quad \text{and} \quad H_i(\alpha_3) = F_i(0, \alpha_3).$$

The integral

$$N(P) = \int_{\mathbb{T}^2} \mathcal{F}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha}$$

gives the number of solutions of the system (8.1) with

$$x_i \in \mathcal{A}(P, R) \ \ (1 \le i \le m+h-2) \quad \text{and} \quad 1 \le x_i \le P \ \ (m+h-1 \le i \le s).$$

We estimate $N(P)$ by dissecting into major and minor arcs as in (7.2). The following lemma takes care of the minor arcs.

Lemma 8.3. *For some $\delta > 0$, one has*

$$\int_{\mathfrak{m}} \mathcal{F}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \ll P^{s-8-\delta}.$$

Proof. By using the trivial inequality $|z_1 \cdots z_r| \le |z_1|^r + \cdots + |z_r|^r$, we find that

$$\int_{\mathfrak{m}} \mathcal{F}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \ll \left(\sup_{\boldsymbol{\alpha} \in \mathfrak{m}} |F_l(\boldsymbol{\alpha})| \right)^2 \int_{\mathbb{T}^2} |f_i(\boldsymbol{\alpha})|^{h-2} |g_j(\alpha_5)|^m |H_k(\alpha_3)|^n \, d\boldsymbol{\alpha}$$

for some i, j, k, and l satisfying

(8.3) $$m+1 \le i \le m+h-2, \quad 1 \le j \le m, \quad m+h+1 \le k \le s,$$

and $l = m+h-1$ or $m+h$. We introduce the abbreviations

(8.4) $$f = |f_i(\boldsymbol{\alpha})|, \qquad g = |g_j(\alpha_5)|, \qquad H = |H_k(\alpha_3)|.$$

Now by arguing as in the proof of Wooley [37], Lemma 7.3, we find that

$$f^{h-2} g^m H^n \ll P^{s-32} \left(f^{30} + f^{16} H^{14} + f^{14} g^{16} + g^{16} H^{14} \right).$$

By combining the results of §6 and [32] with the argument of the proof of [37], Lemma 7.2, we find that

$$\int_{\mathbb{T}^2} \left(f^{30} + f^{16} H^{14} + f^{14} g^{16} + g^{16} H^{14} \right) d\boldsymbol{\alpha} \ll P^{22.113201} + P^{22.077363}.$$

Finally, by applying Theorem 5.1 of Baker [2] as in the proof of Lemma 7.1, we obtain

$$\sup_{\boldsymbol{\alpha} \in \mathfrak{m}} |F_l(\boldsymbol{\alpha})| \ll P^{15/16+\varepsilon},$$

and the lemma now follows on noting that that $1/8 > 0.113201$. $\square$

We now define the pruned major arcs $\mathfrak{N}$ as in (7.7). In order to handle the pruning argument, we need some further mean value estimates.

Lemma 8.4. *Suppose that i, j, and k satisfy (8.3). Then one has the estimates*

$$\int_{\mathbb{T}^2} |f_i(\boldsymbol{\alpha})^{18} H_k(\alpha_3)^8| \, d\boldsymbol{\alpha} \ll P^{18}, \qquad \int_{\mathbb{T}^2} |g_j(\alpha_5)^{18} H_k(\alpha_3)^8| \, d\boldsymbol{\alpha} \ll P^{18},$$

$$\int_{\mathbb{T}^2} |f_i(\boldsymbol{\alpha})^8 g_j(\alpha_5)^{18}| \, d\boldsymbol{\alpha} \ll P^{18}, \qquad \int_{\mathbb{T}^2} |f_i(\boldsymbol{\alpha})|^{32} \, d\boldsymbol{\alpha} \ll P^{24}.$$

Proof. By considering the underlying diophantine equations and applying the Hardy-Littlewood method, it follows easily from the results of [32] that

$$\int_0^1 |g(\alpha)|^{18} \, d\alpha \ll P^{13},$$

and one also knows from Theorem 2 of Vaughan [28] that

$$\int_0^1 |H(\alpha)|^8 \, d\alpha \ll P^5.$$

The first three estimates of the lemma now follow by imitating the proof of Wooley [37], Lemma 9.1, and the final estimate follows from Lemma 7.3 on making a change of variables. $\square$

We are now in a position to complete the pruning argument. As in the proof of Lemma 8.3, we have

$$\int_{\mathfrak{M} \backslash \mathfrak{N}} \mathcal{F}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \ll P^{s-32} \int_{\mathfrak{M} \backslash \mathfrak{N}} F^2(f^{30} + f^{16} H^{14} + f^{14} g^{16} + g^{16} H^{14}) \, d\boldsymbol{\alpha},$$

where f, g, and H are as in (8.4) for some i, j, and k satisfying (8.3), and where we have written $F = |F_l(\boldsymbol{\alpha})|$ for $l = m+h-1$ or $m+h$. On applying Hölder's inequality, we obtain the estimates

$$\int_{\mathfrak{M} \backslash \mathfrak{N}} F^2 f^{30} \, d\boldsymbol{\alpha} \ll \left(\int_{\mathfrak{M} \backslash \mathfrak{N}} F^{32} \, d\boldsymbol{\alpha} \right)^{1/16} \left(\int_{\mathbb{T}^2} f^{32} \, d\boldsymbol{\alpha} \right)^{15/16},$$

$$\int_{\mathfrak{M} \backslash \mathfrak{N}} F^2 f^{16} H^{14} \, d\boldsymbol{\alpha} \ll \left(\int_{\mathfrak{M} \backslash \mathfrak{N}} F^{18} H^{14} \, d\boldsymbol{\alpha} \right)^{1/9} \left(\int_{\mathbb{T}^2} f^{18} H^{14} \, d\boldsymbol{\alpha} \right)^{8/9},$$

$$\int_{\mathfrak{M}\backslash\mathfrak{N}} F^2 g^{16} H^{14}\, d\boldsymbol{\alpha} \ll \left(\int_{\mathfrak{M}\backslash\mathfrak{N}} F^{18} H^{14}\, d\boldsymbol{\alpha}\right)^{1/9} \left(\int_{\mathbb{T}^2} g^{18} H^{14}\, d\boldsymbol{\alpha}\right)^{8/9},$$

$$\int_{\mathfrak{M}\backslash\mathfrak{N}} F^2 f^{14} g^{16}\, d\boldsymbol{\alpha} \ll \left(\int_{\mathfrak{M}\backslash\mathfrak{N}} F^{18} f^{14}\, d\boldsymbol{\alpha}\right)^{1/9} \left(\int_{\mathbb{T}^2} f^{14} g^{18}\, d\boldsymbol{\alpha}\right)^{8/9}.$$

After making trivial estimates and recalling Lemmas 7.2 and 8.4, we find that each of the four integrals above is $O(P^{24} W^{-\sigma})$ for some $\sigma > 0$, and we conclude that

$$(8.5) \qquad \int_{\mathfrak{M}\backslash\mathfrak{N}} \mathcal{F}(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} \ll P^{s-8} W^{-\sigma}.$$

It therefore suffices to deal with the pruned major arcs $\mathfrak{N}$. On taking W to be a suitable power of $\log P$, we find as in §7 that

$$(8.6) \qquad \int_{\mathfrak{N}} \mathcal{F}(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} = \mathfrak{S}(W) J(W) + O(P^{s-8} W^{-\sigma})$$

for some $\sigma > 0$, where $\mathfrak{S}(W)$ is as in (7.11), and where

$$J(W) = \int_{-WP^{-5}}^{WP^{-5}} \int_{-WP^{-3}}^{WP^{-3}} \prod_{i=1}^{m+h-2} w_i(\boldsymbol{\beta}) \prod_{i=m+h-1}^{s} v_i(\boldsymbol{\beta})\, d\boldsymbol{\beta}.$$

The completed singular integral is now given by

$$J = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \prod_{i=1}^{m+h-2} w_i(\boldsymbol{\beta}) \prod_{i=m+h-1}^{s} v_i(\boldsymbol{\beta})\, d\boldsymbol{\beta},$$

and by the appropriate modifications of (7.5) and (7.6), we find that

$$J \ll P^s \int_0^{\infty} \int_0^{\infty} (1+P^3\beta_3)^{-n/5}(1+P^5\beta_5)^{-m/5}(1+P^3\beta_3+P^5\beta_5)^{-h/5}\, d\boldsymbol{\beta}.$$

In view of (8.2), we see that whenever $s \geq 21$ one has

$$J \ll P^s \int_0^{\infty} \int_0^{\infty} (1+P^3\beta_3)^{-6/5}(1+P^5\beta_5)^{-6/5}\, d\boldsymbol{\beta} \ll P^{s-8}$$

and $J - J(W) \ll P^{s-8} W^{-1/5}$. Moreover, the argument of Lemma 7.4 applies with only minor changes to show that $J \gg P^{s-8}$. For the singular series, we first note that by (7.4) we have

$$S(q) \ll \sum_{\substack{a=1 \\ (q,a,b)=1}}^{q} \sum_{b=1}^{q} (q^{-1/5+\varepsilon})^s (q,a)^{m/5} (q,b)^{n/5},$$

since $(q,b)^{1/3}q^{2/3} \leq (q,b)^{1/5}q^{4/5}$. Whenever $T \geq 1$, we have as in the proof of [37], Lemma 10.6, that

$$\sum_{c=1}^{q} (q,c)^T \leq \sum_{d|q} d^T (q/d) \ll q^{T+\varepsilon},$$

so on recalling (8.2), we see that

$$S(q) \ll q^{-s/5} \sum_{a=1}^{q} (q,a)^3 \sum_{b=1}^{q} (q,b) \ll q^{4-s/5+\varepsilon}.$$

Whenever $s \geq 26$, it follows that $\mathfrak{S} \ll 1$ and $\mathfrak{S} - \mathfrak{S}(W) \ll W^{-\tau}$ for some $\tau > 0$. Moreover, the argument of the proof of Lemma 7.5 allows us to conclude that $\mathfrak{S} > 0$. It therefore follows from (8.6) that

$$\int_{\mathfrak{N}} \mathcal{F}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \gg P^{s-8},$$

and the proof of Theorem 1.2 is now completed on recalling (8.5) and Lemma 8.3.

References

[1] R. C. Baker, *Weyl sums and diophantine approximation*, J. London Math. Soc. (2) **25** (1982), 25–34.

[2] _____, *Diophantine inequalities*, Clarendon Press, Oxford, 1986.

[3] _____, *Diagonal cubic equations II*, Acta Arith. **53** (1989), 217–250.

[4] _____, *Correction to 'Weyl sums and diophantine approximation'*, J. London Math. Soc. (2) **46** (1992), 202–204.

[5] R. C. Baker and J. Brüdern, *On pairs of additive cubic equations*, J. reine angew. Math. **391** (1988), 157–180.

[6] B. J. Birch and H. Davenport, *On a theorem of Davenport and Heilbronn*, Acta Math. **100** (1958), 259–279.

[7] Z. I. Borevich and I. R. Shafarevich, *Number theory*, Translated from the Russian by Newcomb Greenleaf, Pure and Applied Mathematics, vol. 20, Academic Press, New York-London, 1966.

[8] J. Brüdern, *A problem in additive number theory*, Math. Proc. Cambridge Philos. Soc. **103** (1988), 27–33.

[9] _____, *On pairs of diagonal cubic forms*, Proc. London Math. Soc. (3) **61** (1990), 273–343.

[10] J. Brüdern and R. J. Cook, *On simultaneous diagonal equations and inequalities*, Acta Arith. **62** (1992), 125–149.

[11] J. Brüdern and T. D. Wooley, *On Waring's problem: Two cubes and seven biquadrates*, Tsukuba J. Math. **24** (2000), 387–417.

[12] R. J. Cook, *Simultaneous quadratic equations*, J. London Math. Soc. (2) **4** (1971), 319–326.

[13] _____, *Pairs of additive equations*, Michigan Math. J. **19** (1972), 325–331.

[14] H. Davenport and D. J. Lewis, *Cubic equations of additive type*, Philos. Trans. Roy. Soc. Ser. A **261** (1966), 97–136.

[15] _____, *Simultaneous equations of additive type*, Philos. Trans. Roy. Soc. Ser. A **264** (1969), 557–595.

[16] M. Dodson, *Homogeneous additive congruences*, Philos. Trans. Roy. Soc. London Ser. A **261** (1967), 163–210.

[17] J. F. Gray, *Diagonal forms of prime degree*, Ph.D. thesis, University of Notre Dame, 1958.

[18] C. Hooley, *On nonary cubic forms*, J. reine angew. Math. **386** (1988), 32–98.

[19] L.-K. Hua, *Additive theory of prime numbers*, A.M.S., Providence, Rhode Island, 1965.

[20] M. P. Knapp, *Diagonal equations of different degrees over p-adic fields* (submitted).

[21] D. J. Lewis, *Cubic forms over algebraic number fields*, Mathematika **4** (1957), 97–101.

[22] S. T. Parsell, *On simultaneous diagonal inequalities*, J. London Math. Soc. (2) **60** (1999), 659–676.

[23] _____, *On simultaneous diagonal inequalities II*, Mathematika (to appear).

[24] S. T. Parsell and T. D. Wooley, *On pairs of diagonal quintic forms*, Compositio Math. (in press).

[25] _____, *A quasi-paucity problem*, Michigan Math. J. (to appear).

[26] W. M. Schmidt, *Simultaneous rational zeros of quadratic forms*, Seminar on Number Theory, Paris 1980-81, Progr. Math., vol. 22, Birkhäuser, Boston, MA, 1982, pp. 281–307.

[27] R. C. Vaughan, *On pairs of additive cubic equations*, Proc. London Math Soc. (3) **34** (1977), 354–364.

[28] _____, *On Waring's problem for cubes*, J. reine angew. Math. **365** (1986), 122–170.

[29] _____, *A new iterative method in Waring's problem*, Acta Math. **162** (1989), 1–71.

[30] _____, *The Hardy-Littlewood method*, 2nd ed., Cambridge University Press, Cambridge, 1997.

[31] R. C. Vaughan and T. D. Wooley, *On Waring's problem: Some refinements*, Proc. London Math. Soc. (3) **63** (1991), 35–68.

[32] _____, *Further improvements in Waring's problem*, Acta Math. **174** (1995), 147–240.

[33] _____, *On a certain nonary cubic form and related equations*, Duke Math. J. **80** (1995), 669–735.

[34] _____, *Further improvements in Waring's problem, IV: Higher powers*, Acta Arith. **94** (2000), 203–285.

[35] T. D. Wooley, *On simultaneous additive equations III*, Mathematika **37** (1990), 85–96.

[36] _____, *On simultaneous additive equations I*, Proc. London Math. Soc. (3) **63** (1991), 1–34.

[37] _____, *On simultaneous additive equations II*, J. reine angew. Math. **419** (1991), 141–198.

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

[39] _____, *Breaking classical convexity in Waring's problem: Sums of cubes and quasi-diagonal behaviour*, Invent. Math. **122** (1995), 421–451.

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

[41] _____, *On simultaneous additive equations IV*, Mathematika **45** (1998), 319–335.

[42] _____, *Sums of three cubes*, Mathematika (in press).

Department of Mathematics, The Pennsylvania State University, McAllister Building, University Park, PA 16802.
  E-mail address: parsell@alum.mit.edu