Exponential Sums and Diophantine Problems

by

Scott T. Parsell

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mathematics)
in The University of Michigan
1999

Doctoral Committee:

Professor Trevor D. Wooley, Chair
Associate Professor Alexander I. Barvinok
Professor Joel N. Bregman
Assistant Professor Kristin E. Lauter
Professor Hugh L. Montgomery

ACKNOWLEDGEMENTS

First of all, I wish to thank my advisor, Trevor Wooley. This work would not have been possible without his patience, advice, and encouragement. I am also grateful for his generous financial support, which allowed me to attend conferences and to spend the winter semesters focused on research.

I also want to acknowledge the roles of Hugh Montgomery and Bob Vaughan in my number theory education at Michigan. I have benefited from several of the advanced courses they have taught, from their numerous seminar presentations, and from the excellent textbooks they have written.

I am grateful to Kristin Lauter, Alexander Barvinok, and Joel Bregman for their willingness to serve on my committee and for their patience in dealing with scheduling changes and thesis revisions. I also thank Doug Richstone for his role in setting up the committee.

I appreciate the help and advice given to me Greg Martin and Joel Wisdom during the early stages of my research. In particular, Greg provided me with many of the LaTeX templates used in this dissertation, and the Student Number Theory seminar, which he established over four years ago, is still thriving today.

The friends I have made at Michigan, especially Bill Correll, David Dorrough, Eric Freeman, James Nuñez, Victor Scharaschkin, and Chris Swanson, deserve recognition for tolerating me during various parts of the past five years. I am also grateful to my sister and to my friends from MIT for staying in touch and providing encouragement along the way.

Finally, I thank my parents for giving me the foundation necessary for this undertaking and for always having unbounded faith in my abilities.

TABLE OF CONTENTS

CHAPTER I

Introduction

1.1 Waring's Problem

Perhaps the most famous result in additive number theory is that every positive integer can be written as the sum of four squares. First stated explicitly in 1621, the theorem survived a claimed proof by Fermat (who died before disclosing it) and the best efforts of Euler before succumbing to a proof by Lagrange in 1770. Many proofs of this result are known today, a representative sampling of which can be found in [25], [43], and [44].

In 1770, shortly before Lagrange announced his theorem, Edward Waring had come forth with a more sweeping conjecture. Let $g(k)$ denote the smallest integer s, if one exists, such that every positive integer n can be represented in the form

$$n = x_1^k + \cdots + x_s^k \tag{1.1}$$

with $x_i \in \mathbb{N} \cup \{0\}$. Waring stated without proof that $g(2) = 4$, $g(3) = 9$, $g(4) = 19$, "and so on," thus implicitly claiming the existence of $g(k)$ for all k. Of course, Lagrange's Theorem establishes the claim that $g(2) = 4$, since integers congruent to 7 modulo 8 cannot be represented as sums of 3 or fewer squares.

In 1909, Hilbert [28] finally demonstrated the existence of $g(k)$, using Lagrange's Theorem as the base of a difficult induction relying on complicated polynomial identities. Because of its inductive nature, Hilbert's argument does not yield any respectable upper bound for $g(k)$. On the other hand, a lower bound showing that $g(k)$ is necessarily quite large is easily obtained by considering the integer

$$n = 2^k \left[\left(\frac{3}{2} \right)^k \right] - 1.$$

Since $n < 3^k$, any representation of n in the form (1.1) must involve only kth powers of 1 and 2. Clearly, the minimal number of terms in such a representation is obtained by taking $[(\frac{3}{2})^k]-1$ powers of 2 and 2^k-1 powers of 1, whence

$$g(k) \geq 2^k + \left[\left(\frac{3}{2}\right)^k\right] - 2. \tag{1.2}$$

Around the time of Hilbert's proof, Wieferich [60] and Kempner [34] were able to show that $g(3)=9$, and it was immediately observed by Landau [35] that only finitely many integers actually require 9 cubes; all others can be represented by 8 or fewer. In fact, Dickson [21] showed in 1939 that 23 and 239 are the only integers that cannot be represented as sums of 8 cubes. In 1943, Linnik [37] took these observations one step further by showing that there are only finitely many integers that cannot be represented as the sum of 7 cubes.

These observations suggest that the enormous size of $g(k)$ results from peculiar difficulties of representing certain small integers. Thus we define a new function $G(k)$ to be the smallest integer s such that every sufficiently large positive integer n can be represented in the form (1.1). While the exact value of $g(k)$ is now known for every k, the problem of determining $G(k)$ turns out to be considerably more difficult. Aside from Lagrange's result that $G(2)=4$, the only value known at present is due to Davenport [16], who showed in 1939 that $G(4)=16$. In particular, our knowledge about sums of cubes is still embarrassingly weak. On combining Linnik's Theorem with an elementary counting argument, one obtains the bounds

$$4 \leq G(3) \leq 7,$$

which remain the best available today, although it is widely conjectured that the lower bound represents the true state of affairs.

1.2 The Hardy-Littlewood Method

In the early 1920's, Hardy and Littlewood [24] devised an analytic approach, known as the circle method, which allows one to derive upper bounds for $G(k)$.

Before describing the method, we briefly introduce some of the notation that will be used throughout. Landau's notation $f(t)=O(g(t))$ means that there exists a positive constant C such that $|f(t)| \leq C|g(t)|$ for all values of t. We will often find it more convenient, however, to write the same statement using Vinogradov's notation, $f(t) \ll g(t)$. If $f(t) \ll g(t)$ and $f(t) \gg g(t)$, then we write $f(t) \asymp g(t)$. Finally, we write $f(t)=o(g(t))$ when $f(t)/g(t) \to 0$ as $t \to \infty$ and $f(t) \sim g(t)$ when $f(t)/g(t) \to 1$ as $t \to \infty$.

Returning now to Waring's problem, we write $P=[n^{1/k}]$, and let

$$R_{s,k}(n) = \mathrm{card}\{\mathbf{x} \in [1,P]^s \cap \mathbb{Z}^s : n = x_1^k + \cdots + x_s^k\}.$$

If we further let $e(z) = e^{2\pi i z}$ and define the exponential sum

$$f(\alpha) = \sum_{1 \le x \le P} e(\alpha x^k), \tag{1.3}$$

then on noting the orthogonality relations

$$\int_0^1 e(\alpha m)\, d\alpha = \begin{cases} 1 & \text{if } m = 0, \\ 0 & \text{if } m \in \mathbb{Z} \setminus \{0\}, \end{cases} \tag{1.4}$$

one sees immediately that

$$R_{s,k}(n) = \int_0^1 f(\alpha)^s e(-\alpha n)\, d\alpha. \tag{1.5}$$

The strategy for evaluating this integral is to dissect the unit interval into major and minor "arcs," with the major arcs consisting of points that are well-approximated by a rational number with small denominator. We should remark that the term "arcs" persists as a result of Hardy and Littlewood's original approach to the problem, in which they used Cauchy's integral formula to represent $R_{s,k}(n)$ as the integral over a circle in the complex plane. The above set-up reflects a later simplification due to Vinogradov [57].

Since there are roughly $n^{s/k}$ total choices for the variables $x_1, \dots, x_s$, a probabilistic argument suggests that one may expect to have $R_{s,k}(n) \asymp n^{s/k-1}$, and in fact this turns out to be the correct order of magnitude provided that s is sufficiently large in terms of k. Roughly speaking, one shows that a contribution of this size arises from the integral over the major arcs and that the integral over the minor arcs is negligible by comparison as $n \to \infty$.

To be more specific, let δ be a small positive number, and define the major arcs by

$$\mathfrak{M} = \bigcup_{\substack{0 \le a \le q \le P^\delta \\ (a,q)=1}} \mathfrak{M}(q,a),$$

where

$$\mathfrak{M}(q,a) = \{\alpha \in [0,1] : |\alpha - a/q| \le P^{\delta - k}\},$$

and write $\mathfrak{m} = [0,1] \setminus \mathfrak{M}$ for the minor arcs. Then one hopes to show that

$$\int_{\mathfrak{M}} f(\alpha)^s e(-\alpha n)\, d\alpha \gg P^{s-k},$$

where the implicit constant may depend on s and k. Hence, if one can also show that

$$\int_{\mathfrak{m}} |f(\alpha)|^s d\alpha = o(P^{s-k}), \tag{1.6}$$

then it will follow that $R_{s,k}(n) > 0$ when n is sufficiently large.

When α is not close to a rational number with small denominator, one expects that the terms in the sum (1.3) will behave somewhat randomly and hence that enough cancellation will occur to establish (1.6). In 1916, Weyl [59] was able to show that for any $\varepsilon > 0$ one has

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

and in 1938 Hua [31] established the mean value estimate

$$\int_0^1 |f(\alpha)|^{2s} d\alpha \ll P^{2s-k+\varepsilon},$$

for $s \geq 2^{k-1}$ by interpreting the integral as the number of solutions of the equation

$$x_1^k + \cdots + x_s^k = y_1^k + \cdots + y_s^k \tag{1.7}$$

with $x_i, y_i \in [1, P] \cap \mathbb{Z}$. Combining these two results, one finds that (1.6) holds whenever $s \geq 2^k + 1$, and this completes the analysis of the minor arcs.

On $\mathfrak{M}(q, a)$, one has $\alpha = a/q + \beta$, where $|\beta| \leq P^{\delta-k}$, so one hopes to relate $f(\alpha)$ to $f(a/q)$. By sorting the sum (1.3) into arithmetic progressions modulo q, we have

$$f(\alpha) = \sum_{r=1}^{q} e\left(\frac{ar^k}{q}\right) \sum_{x \equiv r\ (q)} e(\beta x^k).$$

Since β is small, the function $e(\beta x^k)$ is not oscillating rapidly, so we are able to replace the inner sum by an integral with relatively small error. Hence one obtains the approximation

$$f(\alpha) \sim q^{-1} S(q, a) v(\beta),$$

where

$$S(q, a) = \sum_{1 \leq x \leq q} e\left(\frac{ax^k}{q}\right)$$

and

$$v(\beta) = \int_0^P e(\beta\gamma^k)\, d\gamma.$$

After some analysis, this leads to the factorization

$$\begin{aligned}\int_{\mathfrak{M}} f(\alpha)^s e(-\alpha n)\,d\alpha &\sim \sum_{q\le P^\delta}\sum_{\substack{1\le a\le q\\(a,q)=1}} (q^{-1}S(q,a))^s e(-an/q)\int_{-P^{\delta-k}}^{P^{\delta-k}} v(\beta)^s e(-\beta n)\,d\beta\\ &\sim \mathfrak{S}_{s,k}(n)J_{s,k}(n),\end{aligned}$$

where

$$J_{s,k}(n)=\int_{-\infty}^{\infty} v(\beta)^s e(-\beta n)\,d\beta$$

is known as the singular integral and

$$\mathfrak{S}_{s,k}(n)=\sum_{q=1}^{\infty}\sum_{\substack{1\le a\le q\\(a,q)=1}} (q^{-1}S(q,a))^s e(-an/q)$$

is the singular series. The singular integral captures information about the density of real solutions to (1.1), and in fact it follows from Fourier's Integral Theorem (see for example Davenport [18]) that

$$J_{s,k}(n)\sim \frac{\Gamma(1+1/k)^s}{\Gamma(s/k)}n^{s/k-1}.$$

The positivity of the singular series depends on the p-adic solubility of (1.1), which one establishes by a combination of analytic and Hensel's Lemma-type arguments. Thus when $s\ge 2^k+1$, we obtain the asymptotic formula

$$R_{s,k}(n)\sim \mathfrak{S}_{s,k}(n)\frac{\Gamma(1+1/k)^s}{\Gamma(s/k)}n^{s/k-1},$$

and it follows that $G(k)\le 2^k+1$.

The above result, combined with numerical work on smaller values of n, has led to the determination of $g(k)$ for all values of k. The precise formula is somewhat complicated to state, but the lower bound (1.2) is is fact attained for all but at most finitely many values of k.

Refinements of the above methods have led to substantial improvements in the ensuing bounds for $G(k)$. In particular, Vinogradov's Mean Value Theorem (see for example [55]) allows one to obtain estimates of nearly the same strength as Hua's with s much smaller than 2^k. Moreover, these estimates can be transformed, via the large sieve inequality or similar methods, to yield improved versions of Weyl's inequality. Within this framework of ideas, Vinogradov [58] established the bound

$$G(k)\le 2k\log k(1+o(1))$$

in 1959, but no further improvements of any significance were made over the next 30 years.

1.3 Smooth Numbers and the Iterative Methods

A major breakthrough in the analysis of Waring's problem occurred in 1989 with the work of Vaughan [53]. His contribution was to exploit properties of "smooth" numbers, *i.e.* numbers without large prime factors, to set up an iterative method for estimating the number of solutions of the auxiliary equation (1.7), which arises in the treatment of the minor arcs. Let R be a small power of P, and write

$$\mathcal{A}(P,R) = \{n \in [1,P] \cap \mathbb{Z} : p|n,\ p \text{ prime} \Rightarrow p \leq R\} \tag{1.8}$$

for the set of R-smooth numbers up to P. It is shown, for example in [55], that

$$\operatorname{card}(\mathcal{A}(P,R)) \gg P, \tag{1.9}$$

and hence one may restrict attention to representations of n as sums of kth powers of smooth numbers without serious loss. Moreover, the smooth numbers possess convenient factoring properties, which, roughly speaking, allow one to introduce a congruence condition on some of the variables in (1.7). Thus we define the exponential sum

$$f(\alpha; P, R) = \sum_{x \in \mathcal{A}(P,R)} e(\alpha x^k)$$

and observe that the mean value

$$S_s(P,R) = \int_0^1 |f(\alpha; P, R)|^{2s} d\alpha$$

counts the number of solutions of (1.7) with $x_i, y_i \in \mathcal{A}(P,R)$ and hence in particular is bounded above by the number of solutions of the equation

$$z^k + x_1^k + \cdots + x_{s-1}^k = w^k + y_1^k + \cdots + y_{s-1}^k \tag{1.10}$$

with $1 \leq z, w \leq P$ and $x_i, y_i \in \mathcal{A}(P,R)$. Notice that if $x \in \mathcal{A}(P,R)$ then for any $M > x$ there is a divisor of x lying between M and MR. By applying this observation to the x_i and y_i and then using Hölder's inequality to uniformize the choice of divisor, one is able to relate the number of solutions of (1.10) to the number of solutions of

$$z^k + q^k(u_1^k + \cdots + u_{s-1}^k) = w^k + q^k(v_1^k + \cdots + v_{s-1}^k) \tag{1.11}$$

with

$$1 \leq z, w \leq P, \quad u_i, v_i \in \mathcal{A}(P^{1-\theta}, R), \quad \text{and} \quad P^\theta < q \leq P^\theta R,$$

where θ is a parameter satisfying $\theta \le 1/k$. The congruence condition implicit in (1.11) now allows one to classify solutions according to the common residue class of z^k and w^k modulo q^k, and then by applying Cauchy's inequality to the underlying exponential sums, one reduces to the consideration of solutions in which $z \equiv w \pmod{q^k}$. Hence one may write $w = z + hq^k$ and thus set up a differencing procedure which is "efficient" in the sense that the parameter h is bounded by $P^{1-k\theta}$ instead of P. It is now easy to obtain estimates for $S_s(P, R)$ from estimates for $S_{s-1}(P, R)$ by fixing z, q, and h, and considering the underlying mean values. Thus we have an iterative procedure for estimating $S_s(P, R)$, and this leads to an improved analysis of the minor arcs.

Wooley [62] has refined Vaughan's method to allow one to repeat the differencing process, only fixing z, q, and h trivially after several (perhaps as many as $k-1$) efficient differences have been taken. With this refinement, Wooley was able to halve Vinogradov's bound for $G(k)$, showing in [62] that

$$G(k) \le k(\log k + \log\log k + O(1)). \tag{1.12}$$

We remark that a simple counting argument shows that $G(k) \ge k+1$, and it is expected that this lower bound represents the true state of affairs for most values of k. In certain cases, however, local solubility obstructions may require that $G(k)$ be somewhat larger. Define $\Gamma(k)$ to be the least integer s such that for every n and q the congruence $x_1^k + \cdots + x_s^k \equiv n \pmod{q}$ has a solution with $(x_1, q) = 1$.

Conjecture. *One has* $G(k) = \max(k+1, \Gamma(k))$.

Unfortunately, the technology leading to (1.12) offers little hope of proving such a statement, as the circle method cannot possibly succeed in its present form when $s \le 2k$. In fact, the calculation of $\Gamma(k)$ is a difficult problem in its own right. It is known that $\Gamma(k) \le 4k$ and that this upper bound is attained whenever k is a power of 2, but it is an open question to determine whether $\liminf \Gamma(k) > 3$.

Vaughan and Wooley [56] have introduced iterative schemes that lead to respectable bounds for $G(k)$ for small $k \ge 5$, but even these do not approach what is conjectured. For example, the inequalities

$$6 \le G(5) \le 17, \quad 9 \le G(6) \le 24, \quad 8 \le G(7) \le 33, \quad \text{and} \quad 32 \le G(8) \le 42$$

are the best currently available.

An exposition of the main ideas underlying the use of smooth numbers in additive number theory can be found in Vaughan [54].

1.4 Additive Equations and Inequalities

The Hardy-Littlewood method may also be applied to show that the additive equation

$$c_1x_1^k + \cdots + c_sx_s^k = 0 \tag{1.13}$$

has non-trivial integer solutions when the c_i are nonzero integers (not all of the same sign when k is even) and s is sufficiently large in terms of k. By (1.4), the number of solutions of (1.13) with $x_1, \ldots, x_s \in [1, P] \cap \mathbb{Z}$ is given by

$$N_{s,k}(P) = \int_0^1 \left(\prod_{i=1}^{s} f(c_i\alpha) \right) d\alpha,$$

where $f(\alpha)$ is as in (1.3), and one finds via Hölder's inequality and changes of variable that essentially all of the technology discussed for Waring's problem applies. The only serious difficulty lies in satisfying the p-adic solubility condition and hence proving that the singular series is bounded away from zero. When $k = p - 1$ for some prime p, one must in fact take $s \geq k^2 + 1$, for it is easily seen that the equation

$$(x_1^{p-1} + \cdots + x_{p-1}^{p-1}) + p(y_1^{p-1} + \cdots + y_{p-1}^{p-1}) + \cdots + p^{p-2}(z_1^{p-1} + \cdots + z_{p-1}^{p-1}) = 0$$

has $(p-1)^2$ variables and no non-trivial p-adic solution. It is a classical result of Meyer [40] that $5 = 2^2 + 1$ variables suffice when $k = 2$, and Baker [6] showed in 1989 that 7 variables suffice when $k = 3$. In contrast to the situation for Waring's problem, Baker's result is easily seen to be best possible by considering the equation

$$x_1^3 + 2x_2^3 + 7(x_3^3 + 2x_4^3) + 49(x_5^3 + 2x_6^3) = 0$$

over $\mathbb{Q}_7$. Meyer's result is also seen to be best possible by considering an equation of similar form in 4 variables.

If the integers c_i in (1.13) are replaced by arbitrary real numbers λ_i, then one should not expect to find integers $x_1, \ldots, x_s$ such that $\lambda_1x_1^k + \cdots + \lambda_sx_s^k = 0$, but one can instead ask whether the inequality

$$|\lambda_1x_1^k + \cdots + \lambda_sx_s^k| < \varepsilon \tag{1.14}$$

has non-trivial integer solutions for arbitrarily small ε. In 1946, Davenport and Heilbronn [20] showed that (1.14) has infinitely many integral solutions for any $\varepsilon > 0$ when $k = 2$ and $s \geq 5$, provided that $\lambda_1, \ldots, \lambda_s$ are not all of the same sign. Moreover, their proof easily generalizes to show that $s = 2^k + 1$ variables suffice for general k, again provided that the λ_i are not all of the same sign when k is even. Davenport and Heilbronn established their result by developing an appropriate version of the circle method. While the analytic set-up is now less obvious, the analysis actually turns out to be much easier, owing to the absence of p-adic solubility considerations. We first observe that it suffices to establish the result for $\varepsilon = 1$, since one may then replace λ_i by λ_i/ε. Furthermore, we may assume by relabeling variables that λ_1/λ_2 is irrational, since we may appeal to the theory of additive equations if all the λ_i are in rational ratio. One proceeds by choosing a suitable "kernel" function, for example

$$K(\alpha) = \left(\frac{\sin \pi\alpha}{\pi\alpha}\right)^2, \tag{1.15}$$

whose Fourier transform has properties useful for detecting solutions of (1.14). With the above choice, it is a simple exercise in the calculus of residues to show that

$$\hat{K}(t) = \int_{-\infty}^{\infty} K(\alpha)e(\alpha t)\, d\alpha = \max(0, 1 - |t|),$$

from which it follows that

$$\theta_{s,k}(P) = \int_{-\infty}^{\infty} \left(\prod_{i=1}^{s} f(\lambda_i\alpha)\right) K(\alpha)\, d\alpha$$

is a lower bound for the number of integral solutions of (1.14) with $\mathbf{x} \in [1, P]^s$. Hence it suffices to show that $\theta_{s,k}(P) \to \infty$ along some infinite sequence of P, the sequence here being determined by the denominators of the convergents to the continued fraction for λ_1/λ_2. One now dissects the real line into a single major arc near zero, two minor arcs covering an intermediate range on each side of the major arc, and two large "trivial arcs." The analysis of the minor and trivial arcs requires the use of a mean value estimate such as Hua's Lemma, and the minor arcs additionally require some version of Weyl's inequality. The analysis of the major arc is relatively easy, as one has to deal with a singular integral but no singular series.

There has been considerable interest in quadratic and cubic inequalities in recent years. Most notably, Margulis [39] resolved a conjecture of Oppenheim by proving that any indefinite quadratic form (not necessarily diagonal) in at least 3 variables, takes arbitrarily small values at integer points, provided that not all its coefficients are in rational ratio. This is easily seen to be best possible, since such a statement in two variables would assert that certain algebraic numbers possess rational approximations of higher quality than allowed by Liouville's Theorem (see for example [25]). Recently, Baker, Brüdern, and Wooley [7] showed that 7 variables suffice to solve (1.14) in the cubic case, and they in fact obtained a quantitative version of this result in which ε is replaced by an explicit function of $\mathbf{x}$. It is still possible, however, that as few as 3 variables suffice in the cubic case, and perhaps even for larger k.

A problem of related interest is the so-called "fractional parts" problem, in which we consider inequalities modulo 1. Write $||x||$ for the distance from x to the nearest integer. Then, as a simple example, one tries to establish that for $N > N_0(\varepsilon, k)$

$$\min_{1 \leq n \leq N} ||\alpha n^k|| < N^{\varepsilon - \tau(k)},$$

with $\tau(k)$ as large as possible. The best result on this particular problem is due to Wooley [65], who was able to take $\tau(k)^{-1} \sim k \log k$. Here one does not use the circle method directly, but exponential sum estimates are still the key ingredient in the analysis. When α has major arc-type rational approximations, then the statement is generally easy to prove. Otherwise, one can use some version of Weyl's inequality in combination with a version of the Erdös-Turán inequality (see for example [5], [42]) to bound the discrepancy between the actual and expected distribution of the sequence αn^k modulo 1.

The above methods may be further generalized to investigate systems of equations and inequalities. Let $F_1, \dots, F_t$ be diagonal forms of degree k with real coefficients in s variables, and let ε be a positive real number. The solubility of the system of inequalities

$$|F_1(\mathbf{x})| < \varepsilon, \ \dots \ , |F_t(\mathbf{x})| < \varepsilon \tag{1.16}$$

in integers $x_1, \dots, x_s$ has been considered by a number of authors over the last quarter-century, starting with the work of Cook [15] and Pitman [48] on the case $t = 2$. More recently, Brüdern and Cook [14] have shown that the above system is soluble provided that s is sufficiently large in terms of k and t and that the forms $F_1, \dots, F_t$ satisfy certain additional conditions.

What has not yet been considered is the possibility of allowing the forms $F_1, \dots, F_t$ to have different degrees. However, the recent work of Wooley [61], [72] on the corresponding problem for equations has made the study of such systems a feasible prospect. Our first result takes an initial step in that direction by studying the analogue of the system considered in [61] and [72]. Let $\lambda_1, \dots, \lambda_s$ and $\mu_1, \dots, \mu_s$ be real numbers such that for each i at least one of λ_i or μ_i is nonzero. We define the forms

$$\begin{aligned} F(\mathbf{x}) &= \lambda_1 x_1^3 + \cdots + \lambda_s x_s^3 \\ G(\mathbf{x}) &= \mu_1 x_1^2 + \cdots + \mu_s x_s^2 \end{aligned}$$

and consider the solubility of the system of inequalities

$$\begin{aligned} |F(\mathbf{x})| &< (\max |x_i|)^{-\sigma_1} \\ |G(\mathbf{x})| &< (\max |x_i|)^{-\sigma_2} \end{aligned} \tag{1.17}$$

in rational integers $x_1, \ldots, x_s$. In Chapter 2 we establish the following result by employing a two-dimensional version of the Davenport-Heilbronn method described above.

Theorem 1. *Let $s \geq 13$, and let $\lambda_1, \ldots, \lambda_s$ and $\mu_1, \ldots, \mu_s$ be real numbers such that for some i and j the ratios λ_i/λ_j and μ_i/μ_j are algebraic and irrational. Then the simultaneous inequalities* (1.17) *have infinitely many solutions in rational integers provided that*

(a) *$F(\mathbf{x})$ has at least $s - 4$ variables explicit,*

(b) *$G(\mathbf{x})$ has at least $s - 5$ variables explicit,*

(c) *the system $F(\mathbf{x}) = G(\mathbf{x}) = 0$ has a non-singular real solution, and*

(d) *one has $\sigma_1 + \sigma_2 < \frac{1}{12}$.*

If $\Theta_s(P)$ denotes the number of solutions of (1.17) with $\mathbf{x} \in [1, P]^s \cap \mathbb{Z}^s$, then our analysis in Chapter 2 will in fact show that $\Theta_s(P) \gg P^{s-5-\sigma_1-\sigma_2}$ as $P \to \infty$. We also note that condition (c) implies that the quadratic form G is indefinite, which is plainly a necessary requirement for solubility.

A significant question raised by Theorem 1 is whether the requirement that our forms have some pair of coefficients in algebraic ratio can be shown to be necessary. One suspects that such a condition is not needed, but its removal provides a clear obstruction to the method. When either F or G has a large number of zero coefficients, however, we can exploit results for a single inequality to obtain results that do not require the existence of algebraic irrational coefficient ratios. This is discussed more fully in Chapter 2.

Finally, we remark that the iterative methods developed in Wooley [69] potentially allow one to investigate pairs of inequalities of arbitrary degrees, say k and n, possibly showing solubility in roughly $2k \log k$ variables if $k > n$. However, the mean values of the exponential sums relevant for attacking such a problem have not yet been considered in detail, and obtaining sharp estimates for them is likely to be a formidable task. In principle, one can even consider more general systems of t diagonal inequalities having degrees $k_1, \ldots, k_t$ using the methods of [69] along with the Diophantine approximation results of [5], but this is deferred to later work.

1.5 Multiple Exponential Sums

In contrast to the highly-developed theory of exponential sums in a single variable, multiple exponential sums are not well-understood, and consequently relatively little is known about the Diophantine problems that demand their use. Arkhipov, Karatsuba, and Chubarikov [3] have provided estimates for fairly general d-fold sums, but most of these estimates are not explicit enough to be useful in applications. As a first step toward obtaining sharper general estimates, we concentrate most of our analysis on double sums of the form

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

where R is a small power of P. Write $\mathbb{T}^r$ for the r-dimensional torus. We are particularly interested in estimating the mean value

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

which, by orthogonality, counts the number of solutions of the auxiliary system

$$\sum_{m=1}^{s} (x_m^{k-i} y_m^i - \tilde{x}_m^{k-i} \tilde{y}_m^i) = 0 \quad (0 \leq i \leq k) \tag{1.18}$$

with

$$x_m, y_m, \tilde{x}_m, \tilde{y}_m \in \mathcal{A}(P, R) \quad (1 \leq m \leq s). \tag{1.19}$$

This is done in Chapter 3 by building the apparatus of efficient differencing for polynomials of two variables, so that an iterative method like that of Wooley [69] can be implemented. Thus we are able to derive estimates of the form

$$S_s(P, R) \ll P^{4s - k(k+1) + \Delta_s + \varepsilon}, \tag{1.20}$$

where $\Delta_s \to 0$ as $k \to \infty$. To consider the strength of such an estimate, let $J_s(\mathbf{h})$ denote the number of solutions of the system

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

with (1.19). Then one has

$$J_s(\mathbf{h}) = \int_{\mathbb{T}^{k+1}} |f(\boldsymbol{\alpha}; P, R)|^{2s} e(-\boldsymbol{\alpha} \cdot \mathbf{h}) \, d\boldsymbol{\alpha} \ \leq \ S_s(P, R).$$

On the other hand, upon recalling (1.9) one sees that

$$P^{4s} \ll \sum_{\substack{\mathbf{h} \\ |h_i| \leq sP^k}} J_s(\mathbf{h}) \ll P^{k(k+1)} S_s(P, R),$$

and thus

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

the first term arising from the diagonal solutions of (1.18). Hence the estimate (1.20) becomes nearly best possible as $\Delta_s \to 0$.

The following theorem provides a bound of the shape (1.20) by means of single efficient differencing.

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

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

and let s be a positive integer with $s \geq s_1$. Then for any $\varepsilon > 0$ there exists $\eta = \eta(s,k,\varepsilon)$ such that whenever $R \leq P^\eta$ the estimate (1.20) *holds with*

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

Note for example that if $s \geq 2k^2 \log k$ then we have $\Delta_s \sim k^2 e^{-s/k^2} \leq 1$. Whenever Δ_s has the property that, for every $\varepsilon > 0$, there exists $\eta = \eta(s,k,\varepsilon)$ such that (1.20) holds whenever $R \leq P^\eta$, we say that Δ_s is an admissible exponent.

We remark for comparison that Arkhipov, Karatsuba, and Chubarikov [3] have obtained estimates for the number of solutions of the "complete" system

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

with

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

which lead, via a standard argument, to admissible exponents for (1.18) behaving roughly like $k^3 e^{-s/2k^3}$, so that one must take $s \geq 6k^3 \log k$ in most applications.

In Chapter 3, we obtain the following sharper result as a consequence of repeated efficient differencing, yielding admissible exponents that decay in most cases roughly like $k^2 e^{-3s/2k^2}$.

Theorem 3. *Write $r = \left[\frac{k+1}{2}\right]$, and put*

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

Further, define

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

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

For smaller k, one often needs other methods to obtain reasonable admissible exponents, and this is illustrated for $k = 3$ in Chapter 5.

As was the case with Vinogradov's Mean Value Theorem, our mean value estimates may be transformed into Weyl estimates by using the large sieve inequality. Such estimates will be important to the analysis of the minor arcs in subsequent applications of the Hardy-Littlewood method. For example, we have

Theorem 4. *Define $\mathfrak{m}$ to be the set of $\boldsymbol{\alpha} \in \mathbb{R}^{k+1}$ such that whenever $a_i \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $(a_0, \dots, a_k, q) = 1$ and $|q\alpha_i - a_i| \le P^{1/2-k}R^k$ $(0 \le i \le k)$ one has $q > P^{1/2}R^{k+1}$. Then given $\varepsilon > 0$, there exists $\eta = \eta(\varepsilon, k)$ such that whenever $R \le P^\eta$ one has*

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

where

$$\sigma_1(k)^{-1} \sim \tfrac{28}{3} k^3 \log k \quad \textit{as} \quad k \to \infty.$$

A slightly more general version of this theorem will be proved in Chapter 4, but the above version suffices for most of our applications.

1.6 Applications to Diophantine Problems

Estimates of the type given in Theorem 4 have immediate applications to the problem of obtaining localized bounds for the fractional parts of polynomials in two variables. In particular, we have

Theorem 5. *Given $\boldsymbol{\alpha} \in \mathbb{R}^{k+1}$ and $\varepsilon > 0$, there exists $N_0 = N_0(\varepsilon, k)$ such that whenever $N > N_0$ one has*

$$\min_{1 \le m,n \le N} ||\alpha_0 m^k + \alpha_1 m^{k-1} n + \cdots + \alpha_k n^k|| < N^{\varepsilon - \rho(k)},$$

where

$$\rho(k)^{-1} \sim \tfrac{14}{3} k^3 \log k \quad \textit{as} \quad k \to \infty.$$

It is worth noting that when $\alpha_0 = 0$ or $\alpha_k = 0$ a superior result is obtained by specializing one of the variables and applying Wooley [63], Corollary 1.3, to the resulting polynomial in a single variable. Hence in this case one could take

$$\rho(k)^{-1} \sim 4k^2 \log k$$

in Theorem 5, and so our results are only of interest when both α_0 and α_k are nonzero. A more precise expression for $\rho(k)$ will be given in Chapter 4.

Next we consider a generalization of Waring's problem posed by Arkhipov and Karatsuba [2]. As Waring's problem considers representations of a single integer as a sum of terms of the shape x^k, the simplest form of degree k in a single variable, a natural generalization is to consider simultaneous representations of the shape

$$x_1^{k-j} y_1^j + \cdots + x_s^{k-j} y_s^j = n_j \quad (0 \le j \le k), \tag{1.21}$$

since $x^k, x^{k-1}y, \ldots, y^k$ are the simplest forms of degree k in two variables. By the binomial theorem, we see that this is equivalent to representing the polynomial

$$p(t) = \sum_{j=0}^{k} \binom{k}{j} n_j t^j \tag{1.22}$$

as a sum of s kth powers of linear polynomials. That is, we seek to write

$$p(t) = (x_1 t + y_1)^k + \cdots + (x_s t + y_s)^k \tag{1.23}$$

with $x_i, y_i \in \mathbb{N}$. We remark that the analogous problem over the complex numbers has been considered recently by algebraic geometers (see for example [33], [41]). By exploiting a surprising connection with the theory of partial differential operators, one finds that precisely $s = \left\lceil \frac{k+1}{2} \right\rceil$ terms are required to guarantee a representation of the shape (1.23) for arbitrary polynomials of degree k over $\mathbb{C}$. Working over the integers, however, an elementary counting argument shows that one in fact needs $s \gg k^2$. Obviously, there will be no representations of the shape (1.23) if the relative sizes of the n_j are sufficiently disparate. For example, if some one of the n_j is large, then some one of the x_i or y_i must be large, which in turn forces the other n_j to be large. Thus we will need to impose some conditions in order to obtain a result.

Definition. *For fixed s and k, the polynomial $p(t)$ defined by* (1.22) *is said to be* locally representable *if*

(1) *there exist real numbers P, $\mu_0, \ldots, \mu_k$, and $\delta = \delta(s, k, \boldsymbol{\mu}) > 0$ such that*

$$\left| n_j - P^k \mu_j \right| < \delta P^k \qquad (0 \le j \le k) \tag{1.24}$$

and such that the system

$$\eta_1^{k-j}\xi_1^j + \cdots + \eta_s^{k-j}\xi_s^j = \mu_j \qquad (0 \le j \le k) \tag{1.25}$$

has a non-singular real solution with $0 < \eta_i, \xi_i < 1$*, and*

(2) *the system* (1.21) *has a non-singular* p*-adic solution for all primes* p*.*

Now let $G_1^*(k)$ denote the least integer s such that, whenever the polynomial $p(t)$ given by (1.22) is locally representable and $n_0, \dots, n_k$ are sufficiently large, one has the global representation (1.23) for some natural numbers $x_1, \dots, x_s$ and $y_1, \dots, y_s$.

Theorem 6. *One has*

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

We note that Arkhipov and Karatsuba [2] have previously outlined a program for obtaining bounds of the form $G_1^*(k) \le Ck^2 \log k$ using the theory of multiple exponential sums over a complete interval developed in [3]. Theorem 6 thus gives an explicit asymptotic version of this result, showing that one may take $C \sim 14/3$.

In Chapter 5, we sketch a refined analysis that leads to

Theorem 7. *One has* $G_1^*(3) \le 56$*.*

Our final application concerns the density of rational lines on the hypersurface defined by an additive equation. Let $F(\mathbf{x})$ be a form of degree k in s variables, with integer coefficients. In 1945, Brauer [10] used a diagonalization argument to demonstrate the existence of an m-dimensional linear space on the hypersurface $F(\mathbf{x}) = 0$ over some solvable extension of $\mathbb{Q}$, provided that s is sufficiently large in terms of k and m. By refining Brauer's method, Birch [8] obtained the same result over $\mathbb{Q}$ for odd k in 1957. Unfortunately, the elementary methods of Brauer and Birch do not yield any reasonable estimates for the number of variables required, although explicit astronomical bounds have been given recently by Wooley [70]. For small values of k, somewhat more satisfying results have been obtained by Lewis and Schulze-Pillot [36] and Wooley [67], [68]. Up to this point, however, no estimates have been provided for the density of rational lines on a given hypersurface.

In Chapter 4, we obtain an explicit upper bound for the number of variables required to guarantee the expected density of rational lines on the hypersurface $F(\mathbf{x}) = 0$ in the case when F is an additive form of degree k. Let $c_1, \dots, c_s$ be non-zero integers, and let $L_s(P)$ denote the number of distinct lines of the form $\mathbf{x}t + \mathbf{y}$, with $x_i, y_i \in [-P, P] \cap \mathbb{Z}$, that lie on the hypersurface

$$c_1 z_1^k + \cdots + c_s z_s^k = 0. \tag{1.26}$$

Clearly, $L_s(P)$ is related to the number of solutions of the system of equations

$$c_1x_1^{k-j}y_1^j + \cdots + c_sx_s^{k-j}y_s^j = 0 \quad (0 \le j \le k), \tag{1.27}$$

with $x_i, y_i \in [-P, P] \cap \mathbb{Z}$, so the theory of multiple exponential sums is again applicable. Thus in Chapter 4 we are able to prove

Theorem 8. *Suppose that the system of equations* (1.27) *has a non-singular real solution and a non-singular p-adic solution for all primes p. Then one has*

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

for P sufficiently large, provided that

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

We note that, when s is large in terms of k, the theory of a single additive equation discussed in Section 1.4 shows that the hypersurface defined by (1.26) contains "trivial" lines, corresponding to the case where either $x_i = 0$ or $y_i = 0$ for each i in (1.27). By a trivial estimate, however, the number of such lines is $O(P^s)$. Hence in the situation of Theorem 8 we see that most of the points on (1.26) that lie on lines in fact lie on non-trivial lines.

It transpires that the p-adic solubility conditions imposed in the above theorems need only be checked for finitely many primes p, for we show in Chapter 4 using exponential sums that they do in fact hold whenever $p > p_0(k)$ and $s \ge (k+1)^2$. While p-adic solubility issues were considered in detail by Arkhipov [1] in his work on the Hilbert-Kamke problem, such considerations have largely been ignored in the results stated by Arkhipov and Karatsuba [2] on the multidimensional analogue of Waring's problem. It would therefore be desirable (and possibly quite difficult) to give necessary and sufficient conditions for, or prove unconditionally, the p-adic solubility of the systems (1.21) and (1.27) for small primes.

In the cubic case, we are able to establish a version of Theorem 8 that does not require any local solubility hypotheses. Thus in Chapter 5 we prove

Theorem 9. *Suppose that $k = 3$ and $s \ge 58$. Then for P sufficiently large one has*

$$L_s(P) \gg P^{2s-12}.$$

It is worth noting that higher dimensional analogues of Theorems 6–9 would be accessible with a satisfactory theory of higher dimensional exponential sums. Thus in the analogues of Theorems 8 and 9 we would seek estimates for the density of rational linear spaces (*e.g.* planes) of some dimension $m \ge 2$ that lie on (1.26). Some of the technical apparatus for such a program has been laid by Arkhipov, Karatsuba, and Chubarikov [3] in their treatment of d-fold exponential sums, but a substantial refinement of that analysis would be necessary in order to obtain explicit results. It may be possible to use the generality of their analysis as a model for extending the iterative methods we develop in Chapter 3 to d-fold exponential sums over smooth numbers.

CHAPTER II

Simultaneous Diagonal Inequalities

2.1 Overview

Our main goal in this chapter is the proof of Theorem 1. We begin by recalling some of the notation introduced in Section 1.4. Let $\lambda_1, \ldots, \lambda_s$ and $\mu_1, \ldots, \mu_s$ be real numbers such that for each i at least one of λ_i or μ_i is nonzero, and define the forms

$$\begin{aligned} F(\mathbf{x}) &= \lambda_1 x_1^3 + \cdots + \lambda_s x_s^3 \\ G(\mathbf{x}) &= \mu_1 x_1^2 + \cdots + \mu_s x_s^2. \end{aligned}$$

Further, let $\Theta_s(P)$ denote the number of solutions of the system of inequalities

$$\begin{aligned} |F(\mathbf{x})| &< (\max |x_i|)^{-\sigma_1} \\ |G(\mathbf{x})| &< (\max |x_i|)^{-\sigma_2} \end{aligned} \tag{2.1}$$

with $x_1, \ldots, x_s \in [1, P] \cap \mathbb{Z}$. The following is a quantitative version of Theorem 1.

Theorem 2.1. *Let $s \geq 13$, and let $\lambda_1, \ldots, \lambda_s$ and $\mu_1, \ldots, \mu_s$ be real numbers such that for some i and j the ratios λ_i/λ_j and μ_i/μ_j are algebraic and irrational. Further, suppose that*

(a) *$F(\mathbf{x})$ has at least $s-4$ variables explicit,*

(b) *$G(\mathbf{x})$ has at least $s-5$ variables explicit,*

(c) *the system $F(\mathbf{x}) = G(\mathbf{x}) = 0$ has a non-singular real solution, and*

(d) *one has $\sigma_1 + \sigma_2 < \frac{1}{12}$.*

Then one has $\Theta_s(P) \gg P^{s-5-\sigma_1-\sigma_2}$ for P sufficiently large.

When either F or G has a large number of zero coefficients, we can exploit results for a single inequality to obtain a result in which the conditions on the coefficient ratios are not needed.

Theorem 2.2. *Let $\lambda_1, \ldots, \lambda_s$ and $\mu_1, \ldots, \mu_s$ be real numbers. The simultaneous inequalities* (2.1) *have infinitely many solutions in rational integers provided that*

(a) *$F(\mathbf{x})$ has at least 7 variables explicit,*

(b) *$G(\mathbf{x})$ has at least 5 variables explicit,*

(c) *the system $F(\mathbf{x}) = G(\mathbf{x}) = 0$ has a non-singular real solution, and*

(d) *one of the following holds:*

  (i) *at least 4 of the λ_i are zero and* $\max(\sigma_1, \sigma_2) \leq 10^{-5}$*, or*

  (ii) *at least 7 of the μ_i are zero and $\sigma_1 \leq 10^{-4}$.*

We remark that condition (b) is not actually needed to prove the stated version of Theorem 2.2; however, the condition arises naturally in discussing possible improvements on condition (d)(ii), so we state it for convenience.

In Section 2.2, we deduce Theorem 2.2 in an elementary manner from results on a single Diophantine inequality. We also consider a refinement of condition (d)(ii) that would follow from improvements in our understanding of cubic inequalities.

We then prove Theorem 2.1 in Sections 2.3, 2.4, and 2.5, using a two-dimensional version of the Davenport-Heilbronn method. We show that

$$\Theta_s(P) \gg \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{H}(\boldsymbol{\alpha})\mathcal{K}(\boldsymbol{\alpha})\, d\boldsymbol{\alpha},$$

where $\mathcal{H}(\boldsymbol{\alpha})$ is a suitable product of exponential sums (many of which we restrict to smooth numbers) and $\mathcal{K}(\boldsymbol{\alpha})$ is a product of two kernels similar to (1.15). We then dissect the plane in analogy with the one-dimensional Davenport-Heilbronn dissection discussed in Section 1.4. The success of our minor arc analysis depends heavily on an estimate of Wooley [72] for the 10th moment of a certain exponential sum over smooth numbers and also on a result of R. Baker [5] relating the size of a certain exponential sum to the existence of good rational approximations to the coefficients of its argument. The treatment of the major arc is essentially straightforward using the ideas of Wooley [61].

Finally, in Section 2.6, we discuss the possibility of weakening some of the hypotheses imposed in Theorems 2.1 and 2.2.

Throughout our analysis, implicit constants in the notations of Vinogradov and Landau may depend on the coefficients $\lambda_1, \ldots, \lambda_s$ and $\mu_1, \ldots, \mu_s$, the exponents σ_1 and σ_2, and also on any parameters denoted by ε or δ.

The material of this chapter appears in the author's forthcoming publication [47].

2.2 Forms with Many Zero Coefficients

Here we prove Theorem 2.2 using results on a single inequality. We first consider the case (d)(i). The argument is similar to that given in Lemmata 6.3, 6.4, and 6.5 of Wooley [61], but it also incorporates the recent work of Baker, Brüdern, and Wooley [7] on cubic inequalities in 7 variables and makes use of a result of Birch and Davenport [9] on small solutions of quadratic inequalities in 5 variables. We start with an analogue of [61], Lemma 6.3.

Lemma 2.2.1. *Suppose that there is a rearrangement of the variables $x_1, \ldots, x_s$ such that $\lambda_i = 0$ for $i = 1, \ldots, 4$ and $\mu_1, \ldots, \mu_4$ are not all of the same sign. Then Theorem* 2.2 *holds in the case* (d)(i)*.*

Proof. Let $\sigma = 1.43 \times 10^{-4}$ and $\delta = \frac{1}{10}\sigma$. It is easily seen that the main theorem of [7] holds with the above value of σ, although the result is stated with a slightly smaller exponent. Thus by condition (a) of Theorem 2.2, there exist infinitely many $(s-4)$-tuples of integers $(a_5, \ldots, a_s)$ such that

$$\left|\lambda_5 a_5^3 + \cdots + \lambda_s a_s^3\right| < (\max |a_i|)^{-\sigma}. \tag{2.2}$$

Now put $M_i = \mu_i$ for $i = 1, \ldots, 4$, and put

$$M_5 = \mu_5 a_5^2 + \cdots + \mu_s a_s^2.$$

If $|M_5| < (\max |a_i|)^{-\delta}$, then we can take $x_1 = \cdots = x_4 = 0$ and $x_i = a_i$ for $i = 5, \ldots, s$. Otherwise, by the main theorem of [9] we can find (for $\max |a_i|$ sufficiently large) integers $u_1, \ldots, u_5$, not all zero, such that

$$\left|M_1 u_1^2 + \cdots + M_5 u_5^2\right| < (\max |a_i|)^{-\delta} \tag{2.3}$$

and

$$|M_1 u_1^2| + \cdots + |M_5 u_5^2| \ll (\max |a_i|)^{\delta(4+5\delta)} |M_1 \cdots M_5|^{1+\delta}.$$

But $M_5 \ll (\max |a_i|)^2$, so that

$$|u_j| \ll (\max |a_i|)^{1+\frac{\delta}{2}(6+5\delta)} \quad (j = 1, \ldots, 4)$$

and

$$|u_5| \ll (\max |a_i|)^{\frac{\delta}{2}(6+5\delta)}.$$

Hence on putting $\mathbf{x} = (u_1, \ldots, u_4, u_5 a_5, \ldots, u_5 a_s)$, we have

$$\max |x_i| \ll (\max |a_i|)^{1+\frac{\delta}{2}(6+5\delta)}$$

and

$$|F(\mathbf{x})| < |u_5|^3(\max|a_i|)^{-\sigma} \ll (\max|a_i|)^{\frac{3\delta}{2}(6+5\delta)-\sigma}.$$

Thus on taking

$$\varepsilon < \frac{2\sigma - 3\delta(6+5\delta)}{2+\delta(6+5\delta)}$$

we see that for $\max|a_i|$ sufficiently large one has

$$|F(\mathbf{x})| < (\max|x_i|)^{-\varepsilon},$$

and so we may take $\sigma_1 = 1.429 \times 10^{-5}$. Moreover, on taking

$$\gamma < \frac{2\delta}{2+\delta(6+5\delta)}$$

we have

$$|G(\mathbf{x})| < (\max|a_i|)^{-\delta} < (\max|x_i|)^{-\gamma}$$

for $\max|a_i|$ sufficiently large, so we may take $\sigma_2 = 1.429 \times 10^{-5}$. $\square$

When the hypothesis of Lemma 2.2.1 is not satisfied, we need some additional control over the solution to our cubic inequality (2.2) in order to guarantee that the quadratic in (2.3) is indefinite. Specifically, we require the following analogue of [61], Lemma 6.4.

Lemma 2.2.2. *Let $\lambda_1, \ldots, \lambda_t$ $(t \geq 7)$ be non-zero real numbers, and suppose that $(\eta_1, \ldots, \eta_t)$ is a real solution of the equation*

$$\lambda_1 x_1^3 + \cdots + \lambda_t x_t^3 = 0$$

with $0 < \eta_i < 1$ for all i. Then for any $\alpha \in (0,1)$ and $P > P_0(\boldsymbol{\eta}, \boldsymbol{\lambda}, \alpha)$, there exist integers $y_1, \ldots, y_t$ such that

$$|\lambda_1 y_1^3 + \cdots + \lambda_t y_t^3| < (\max|y_i|)^{-\sigma},$$

where $\sigma = 1.43 \times 10^{-4}$ and

$$(1-\alpha)\eta_i P < y_i \leq (1+\alpha)\eta_i P \quad (i = 1, \ldots, t). \tag{2.4}$$

Proof. If the λ_i are all in rational ratio, then the result follows from Lemma 6.4 of [61]. Otherwise, we follow through the analysis of [7], restricting the ranges of summation on the generating functions so that only values of the variables satisfying (2.4) are included. All of the required estimates continue to hold, with only the major arc analysis requiring a slight modification. □

Now we can complete the proof of case (d)(i) by arguing as in the proof of [61], Lemma 6.5. Suppose that at least 4 of the λ_i are zero, and rearrange variables so that $\lambda_1, \dots, \lambda_t \neq 0$ and $\lambda_i = 0$ for $i = t+1, \dots, s$. By condition (c) and the argument of [61], Lemma 6.2, we may assume that the equations $F(\mathbf{x}) = G(\mathbf{x}) = 0$ have a real solution $(\eta_1, \dots, \eta_s)$ with all of the η_i non-zero, and then on replacing λ_i by $-\lambda_i$ if necessary and using homogeneity we may assume that $0 < \eta_i < \frac{1}{2}$ for all i. Further, by Lemma 2.2.1, we may assume that $\mu_{t+1}, \dots, \mu_s$ are all positive, so that

$$\mu_1\eta_1^2 + \cdots + \mu_t\eta_t^2 = -(\mu_{t+1}\eta_{t+1}^2 + \cdots + \mu_s\eta_s^2) = -C < 0.$$

Let α, P, and $(y_1, \dots, y_t)$ be as in Lemma 2.2.2 with

$$\alpha < \frac{2C}{3t}\left(\max|\mu_i|\right)^{-1},$$

and put $M = \mu_1 y_1^2 + \cdots + \mu_t y_t^2$. Then

$$|M + CP^2| \leq P^2(\alpha^2 + 2\alpha)\sum_{i=1}^{t}|\mu_i\eta_i^2| < \frac{1}{2}CP^2,$$

so that

$$M < -\frac{1}{2}CP^2 < 0.$$

Now let $\delta = 1.43 \times 10^{-5}$ as before. If $|M| < P^{-\delta}$, then we can take $x_i = y_i$ for $i = 1, \dots, t$ and $x_{t+1} = \cdots = x_s = 0$. Otherwise, for P sufficiently large, we may use the result of [9] as in the proof of Lemma 2.2.1 to find integers $v_t, \dots, v_s$, not all zero, with

$$|v_t| \ll P^{\frac{\delta}{2}(6+5\delta)} \quad \text{and} \quad |v_i| \ll P^{1+\frac{\delta}{2}(6+5\delta)} \quad (i = t+1, \dots, s)$$

such that

$$|Mv_t^2 + \mu_{t+1}v_{t+1}^2 + \cdots + \mu_s v_s^2| < P^{-\delta}.$$

Proceeding exactly as in the proof of Lemma 2.2.1, we find that

$$\mathbf{x} = (y_1 v_t, \dots, y_t v_t, v_{t+1}, \dots, v_s)$$

satisfies (2.1) with $\sigma_1 = \sigma_2 = 10^{-5}$, and this completes the proof of Theorem 2.2 in the case (d)(i).

The case (d)(ii) of Theorem 2.2 follows immediately from the results of [7], and this completes the proof of the theorem.

We now investigate the possibility of reducing the number of zero coefficients required by condition (d)(ii) from 7 to 6, in accordance with [61] and [72]. Brüdern [13], improving on a result of Pitman and Ridout [49], has shown that if $\lambda_1, \ldots, \lambda_9$ are real numbers with $|\lambda_i| \geq 1$ for all i then there exist integers $x_1, \ldots, x_9$ satisfying

$$|\lambda_1 x_1^3 + \cdots + \lambda_9 x_9^3| < 1$$

and

$$0 < \sum_{i=1}^{9} |\lambda_i x_i^3| \ll_\delta |\lambda_1 \cdots \lambda_9|^{1+\delta}. \tag{2.5}$$

Unfortunately, in order to use this result in an argument like the one in Lemma 2.2.1 we would have to assume that $G(\mathbf{x})$ had at least eight zero coefficients, and in this situation we would do better to apply the results of [12]. Suppose, however, that the above result held with 7 variables instead of 9. Then condition (d)(ii) of Theorem 2.2 could be replaced by

$$\text{(d)(ii)}' \quad \textit{at least } 6 \textit{ of the } \mu_i \textit{ are zero and } \max(\sigma_1, \sigma_2) \leq 10^{-2}.$$

The argument resembles the one above, but an argument like the one ensuing from Lemma 2.2.2 will not be necessary since the quadratic under consideration there will be replaced by a cubic.

Proceeding just as in Lemma 2.2.1, we fix $\sigma < 1/10$ and $\delta = 1/70$. After rearranging variables, we may assume that $\mu_1 = \cdots = \mu_6 = 0$. Now by condition (b) of Theorem 2.2 and an easily obtained quantitative version of the classical Davenport-Heilbronn Theorem, we see that there exist infinitely many $(s-6)$-tuples of integers $(a_7, \ldots, a_s)$ such that

$$\left|\mu_7 a_7^2 + \cdots + \mu_s a_s^2\right| < (\max|a_i|)^{-\sigma}.$$

Now put $\Lambda_i = \lambda_i$ for $i = 1, \ldots, 6$, and put

$$\Lambda_7 = \lambda_7 a_7^3 + \cdots + \lambda_s a_s^3.$$

If $|\Lambda_7| < (\max|a_i|)^{-\delta}$, then we can take $x_1 = \cdots = x_6 = 0$ and $x_i = a_i$ for $i = 7, \ldots, s$. Otherwise, by our hypothesis, we can find (for $\max|a_i|$ sufficiently large) integers $u_1, \ldots, u_7$, not all zero, such that

$$\left|\Lambda_1 u_1^3 + \cdots + \Lambda_7 u_7^3\right| < (\max|a_i|)^{-\delta}$$

and

$$|\Lambda_1 u_1^3| + \cdots + |\Lambda_7 u_7^3| \ll (\max|a_i|)^{\delta(6+7\delta)} |\Lambda_1 \cdots \Lambda_7|^{1+\delta}.$$

But $\Lambda_7 \ll (\max|a_i|)^3$, so that

$$|u_j| \ll (\max|a_i|)^{1+\frac{\delta}{3}(9+7\delta)} \quad (j = 1, \ldots, 6)$$

and

$$|u_7| \ll (\max|a_i|)^{\frac{\delta}{3}(9+7\delta)}.$$

Hence on putting $\mathbf{x} = (u_1, \ldots, u_6, u_7 a_7, \ldots, u_7 a_s)$, we have

$$\max|x_i| \ll (\max|a_i|)^{1+\frac{\delta}{3}(9+7\delta)},$$

so on taking

$$\gamma < \frac{3\delta}{3 + \delta(9 + 7\delta)}$$

we have

$$|F(\mathbf{x})| < (\max|a_i|)^{-\delta} < (\max|x_i|)^{-\gamma}.$$

Furthermore, if

$$\varepsilon < \frac{3\sigma - 2\delta(9 + 7\delta)}{3 + \delta(9 + 7\delta)}$$

then we have

$$|G(\mathbf{x})| < |u_7|^2 (\max|a_i|)^{-\sigma} \ll (\max|a_i|)^{\frac{2\delta}{3}(9+7\delta)2-\sigma},$$

whence for $\max|a_i|$ sufficiently large

$$|G(\mathbf{x})| < (\max|x_i|)^{-\varepsilon}.$$

Thus we may take $\sigma_1 = \sigma_2 = 1.2 \times 10^{-2}$.

We note that throughout our arguments there is some freedom in the choice of the parameter δ, and we have generally chosen it so as to give roughly the same permissible values for σ_1 and σ_2. If so desired, one can alter δ in favor of one exponent or the other and in fact obtain a region of permissible values similar in shape to (but smaller than) the region in Theorem 2.1(d). We do not pursue this refinement here.

2.3 The Davenport-Heilbronn Method

We now set up a two-dimensional version of the Davenport-Heilbronn method, which we will use to prove Theorem 2.1. We may assume (after rearranging variables) that the the first m of the μ_i are zero, that the last n of the λ_i are zero, and that the remaining $h = s - m - n$ indices have both λ_i and μ_i nonzero. Then when $s \geq 13$ we have by conditions (a) and (b) of Theorem 2.1 that

$$0 \leq m \leq 5, \quad 0 \leq n \leq 4, \quad \text{and} \quad h \geq 4. \tag{2.6}$$

Furthermore, we may suppose that λ_I/λ_J and μ_I/μ_J are algebraic irrationals, where

$$I = m + h - 2, \quad J = m + h - 1, \quad \text{and} \quad K = m + h.$$

Let ε be a small positive number, and choose $\eta > 0$ sufficiently small in terms of ε. Take P to be a large positive number, put $R = P^\eta$, and let

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

Write $\boldsymbol{\alpha} = (\alpha, \beta)$, and define generating functions

$$F_i(\boldsymbol{\alpha}) = \sum_{1 \leq x \leq P} e(\lambda_i \alpha x^3 + \mu_i \beta x^2) \tag{2.7}$$

and

$$f_i(\boldsymbol{\alpha}) = \sum_{x \in \mathcal{A}(P,R)} e(\lambda_i \alpha x^3 + \mu_i \beta x^2). \tag{2.8}$$

It will also be convenient to write

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

According to Davenport [17], for every integer r there exists a real-valued even kernel function K of one real variable such that

$$K(\alpha) \ll \min(1, |\alpha|^{-r}) \tag{2.9}$$

and

$$\int_{-\infty}^{\infty} e(\alpha t) K(\alpha)\, d\alpha \begin{cases} = 0, & \text{if } |t| \geq 1, \\ \in [0, 1], & \text{if } |t| \leq 1, \\ = 1, & \text{if } |t| \leq \frac{1}{3}. \end{cases} \tag{2.10}$$

We set

$$\mathcal{K}(\boldsymbol{\alpha}) = K(\alpha P^{-\sigma_1})K(\beta P^{-\sigma_2}).$$

Now let $N(P)$ be the number of solutions of (2.1) with

$$x_i \in \mathcal{A}(P, R) \qquad (i = 1, \ldots, m + h - 3)$$

and

$$1 \leq x_i \leq P \qquad (i = m + h - 2, \ldots, s).$$

By a familiar argument, $N(P)$ is bounded below by $P^{-\sigma_1-\sigma_2}R(P)$, where

$$R(P) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{F}(\boldsymbol{\alpha})\mathcal{G}(\boldsymbol{\alpha})\mathcal{H}(\boldsymbol{\alpha})\mathcal{K}(\boldsymbol{\alpha})\, d\boldsymbol{\alpha}, \tag{2.11}$$

$$\mathcal{F}(\boldsymbol{\alpha}) = \prod_{i=1}^{m+h-3} f_i(\boldsymbol{\alpha}), \quad \mathcal{H}(\boldsymbol{\alpha}) = \prod_{i=m+h-2}^{m+h} F_i(\boldsymbol{\alpha}), \quad \text{and} \quad \mathcal{G}(\boldsymbol{\alpha}) = \prod_{i=m+h+1}^{s} F_i(\boldsymbol{\alpha}).$$

We dissect the plane into three main regions, imitating the standard dissection of the real line used in the treatment of a single inequality. The trivial region is defined by

$$\mathfrak{t} = \{\boldsymbol{\alpha} : |\alpha| > P^{\sigma_1+\varepsilon} \text{ or } |\beta| > P^{\sigma_2+\varepsilon}\}, \tag{2.12}$$

the major arc by

$$\mathfrak{M} = \{\boldsymbol{\alpha} : |\alpha| \leq P^{-9/4} \text{ and } |\beta| \leq P^{-5/4}\}, \tag{2.13}$$

and the minor arcs by

$$\mathfrak{m} = \mathbb{R}^2 \setminus (\mathfrak{t} \cup \mathfrak{M}). \tag{2.14}$$

Our plan is to show that $R(P) \gg P^{s-5}$, with the main contribution coming from the major arc. For r sufficiently large in terms of ε, it follows easily from (2.9) and (2.12) that the contribution to $R(P)$ from the trivial region is $o(P^{s-5})$. In the next section, we consider a finer dissection of the minor arcs, which allows us to show that their contribution to $R(P)$ is also $o(P^{s-5})$, provided that σ_1 and σ_2 are confined to the region specified in Theorem 2.1. Finally, in Section 2.5, we apply standard methods to deal with the major arc.

2.4 The Minor Arcs

We begin by bounding the integral (2.11) in terms of others having somewhat more standard forms. We start by choosing a finite covering of $\mathfrak{m}$ by unit squares of the form $[c, c+1] \times [d, d+1]$. For $\mathfrak{n} \subset \mathfrak{m}$, let $\mathcal{U}_\mathfrak{n}$ denote the square for which the integral

$$\iint\limits_{\mathfrak{n}\cap\mathcal{U}_\mathfrak{n}} |\mathcal{F}(\boldsymbol{\alpha})\mathcal{G}(\boldsymbol{\alpha})\mathcal{H}(\boldsymbol{\alpha})|\, d\boldsymbol{\alpha}$$

is maximal, and write $\mathfrak{n}^* = \mathfrak{n} \cap \mathcal{U}_\mathfrak{n}$. Then for $r > 1$ it follows from (2.9) that

$$\iint\limits_{\mathfrak{n}} |\mathcal{F}(\boldsymbol{\alpha})\mathcal{G}(\boldsymbol{\alpha})\mathcal{H}(\boldsymbol{\alpha})\mathcal{K}(\boldsymbol{\alpha})|\, d\boldsymbol{\alpha} \ll P^{\sigma_1+\sigma_2} \iint\limits_{\mathfrak{n}^*} |\mathcal{F}(\boldsymbol{\alpha})\mathcal{G}(\boldsymbol{\alpha})\mathcal{H}(\boldsymbol{\alpha})|\, d\boldsymbol{\alpha}. \tag{2.15}$$

Furthermore, by arguing as in the proof of Lemma 7.3 of Wooley [61], we see that

$$\iint\limits_{\mathfrak{n}^*} |\mathcal{F}(\boldsymbol{\alpha})\mathcal{G}(\boldsymbol{\alpha})|\, d\boldsymbol{\alpha} \ll \iint\limits_{\mathfrak{n}^*} |f_i(\boldsymbol{\alpha})|^{h-3} |g_j(\alpha)|^m |H_k(\beta)|^n\, d\boldsymbol{\alpha} \tag{2.16}$$

for some i, j, and k (depending on $\mathfrak{n}$) satisfying

$$m+1 \le i \le m+h, \quad 1 \le j \le m, \quad \text{and} \quad m+h+1 \le k \le s.$$

In the course of an argument in which $\mathfrak{n}$ is fixed, we will employ the abbreviations

$$f = |f_i(\boldsymbol{\alpha})|, \quad g = |g_j(\alpha)|, \quad \text{and} \quad H = |H_k(\beta)|.$$

Finally, on recalling (2.6) and again mimicking the arguments of [61], we obtain

$$f^{h-3} g^m H^n \ll P^{s-13}\left(f^{10} + f^u H^{10-u} + g^u H^{10-u} + f^{10-u} g^u\right) \tag{2.17}$$

whenever $5 \le u \le 6$. For convenience, we introduce the notation

$$Q = P^{s-13+\sigma_1+\sigma_2}. \tag{2.18}$$

We are now in a position to make use of certain mean value estimates developed in Wooley [61], [72]. Those that we need are recorded for reference in the following lemma.

Lemma 2.4.1. *Suppose that*

$$m+1 \le i \le m+h, \quad 1 \le j \le m, \quad \textit{and} \quad m+h+1 \le k \le s.$$

Then for any unit square $\mathcal{U} = [c, c+1] \times [d, d+1]$, *we have*

(i) $\iint_{\mathcal{U}} |f_i(\boldsymbol{\alpha})|^{10} \, d\boldsymbol{\alpha} \ll P^{17/3+\varepsilon}$,

(ii) $\iint_{\mathcal{U}} |f_i(\boldsymbol{\alpha})|^{6} |H_k(\beta)|^4 \, d\boldsymbol{\alpha} \ll P^{21/4+\varepsilon}$,

(iii) $\iint_{\mathcal{U}} |g_j(\alpha)|^{6} |H_k(\beta)|^4 \, d\boldsymbol{\alpha} \ll P^{21/4+\varepsilon}$,

(iv) $\iint_{\mathcal{U}} |f_i(\boldsymbol{\alpha})|^{4} |g_j(\alpha)|^6 \, d\boldsymbol{\alpha} \ll P^{21/4+\varepsilon}$,

(v) $\iint_{\mathcal{U}} |f_i(\boldsymbol{\alpha})|^{14} \, d\boldsymbol{\alpha} \ll P^{9}$,

(vi) $\iint_{\mathcal{U}} |f_i(\boldsymbol{\alpha})|^{8} |H_k(\beta)|^5 \, d\boldsymbol{\alpha} \ll P^{8}$,

(vii) $\iint_{\mathcal{U}} |g_j(\alpha)|^{8} |H_k(\beta)|^5 \, d\boldsymbol{\alpha} \ll P^{8}$,

(viii) $\iint_{\mathcal{U}} |f_i(\boldsymbol{\alpha})|^{6} |g_j(\alpha)|^8 \, d\boldsymbol{\alpha} \ll P^{9}$.

Proof. Part (i) follows from Theorem 2 of Wooley [72] on considering the underlying Diophantine equations and making a change of variables. Parts (iii), (v), and (vii) follow from the corresponding parts of Lemmata 7.2, 9.1, and 9.4 of Wooley [61] on making a change of variables and noting that the additional restrictions imposed on the variable ranges in that paper can be removed without affecting the arguments. For the remaining parts, we use the idea of the proof of Lemma 9.1(i) of [61] in a manner typified by (ii): Write

$$s_m(\mathbf{x}, \mathbf{y}) = (x_1^m - y_1^m) + (x_2^m - y_2^m) + (x_3^m - y_3^m)$$

and

$$H(\beta) = \sum_{1 \le x \le P} e(\beta x^2).$$

Then on making the change of variables $\alpha' = \lambda_i \alpha$ and $\beta' = \mu_k \beta$ we have

$$\iint_{\mathcal{U}} |f_i(\boldsymbol{\alpha})|^6 |H_k(\beta)|^4 \, d\boldsymbol{\alpha} \ll \iint_{\mathcal{U}'} \sum_{\mathbf{x},\mathbf{y}} e\left(s_3(\mathbf{x},\mathbf{y})\alpha + \frac{\mu_i}{\mu_k} s_2(\mathbf{x},\mathbf{y})\beta \right) |H(\beta)|^4 \, d\alpha \, d\beta,$$

where the summation is over $\mathbf{x}$ and $\mathbf{y}$ with $x_i, y_i \in \mathcal{A}(P,R)$ and where $\mathcal{U}' = [m_3, n_3] \times [m_2, n_2]$ for some integers m_j and n_j with $n_j - m_j \ll 1$. If we now let

$$c(\mathbf{x},\mathbf{y}) = e\left(\frac{\mu_i}{\mu_k} s_2(\mathbf{x},\mathbf{y})\beta \right),$$

then since $c(\mathbf{x},\mathbf{y})$ is unimodular we obtain

$$\begin{aligned}\iint_{\mathcal{U}} |f_i(\boldsymbol{\alpha})|^6 |H_k(\beta)|^4 \, d\boldsymbol{\alpha} &\ll \int_{m_2}^{n_2} \left(\sum_{\mathbf{x},\mathbf{y}} c(\mathbf{x},\mathbf{y}) \int_{m_3}^{n_3} e(s_3(\mathbf{x},\mathbf{y})\alpha) d\alpha \right) |H(\beta)|^4 \, d\beta \\ &\ll P^{13/4+\varepsilon} \int_0^1 |H(\beta)|^4 \, d\beta \ll P^{21/4+\varepsilon}\end{aligned}$$

on using Theorem 4.4 of Vaughan [53] and considering the underlying Diophantine equations. □

Lemma 2.4.1 allows us to handle regions of $\mathfrak{m}$ on which $\mathcal{H}$ is suitably bounded. Fortunately, when F_I, F_J, or F_K is large, we also obtain a great deal of information from a theorem of Baker [5], a special case of which is recorded below.

Lemma 2.4.2. *Let $P > P_0(\varepsilon)$ and $A > P^{3/4+\varepsilon}$. If $|F_i(\boldsymbol{\alpha})| \geq A$ for some $i = I, J$, or K, then there exists a natural number $q < P^{3+\varepsilon}A^{-3}$ and integers a and b with $(q, a, b) = 1$ such that $|\lambda_i \alpha q - a| < P^{\varepsilon} A^{-3}$ and $|\mu_i \beta q - b| < P^{1+\varepsilon} A^{-3}$.*

Proof. This is Theorem 5.1 of [5] with $T = P^{3/4+\varepsilon}$, $M = 1$, and $k = 3$. □

Lemma 2.4.2 suggests further dissecting $\mathfrak{m}$ according to the behavior of F_I, F_J, and F_K. Thus we start by defining

$$\mathfrak{e} = \{\boldsymbol{\alpha} \in \mathfrak{m} : |F_i(\boldsymbol{\alpha})| \leq P^{3/4+\varepsilon} \text{ for } i = I, J, K\}.$$

Now let

$$\mathfrak{f}(I) = \{\boldsymbol{\alpha} \in \mathfrak{m} : |F_I(\boldsymbol{\alpha})| > P^{3/4+\varepsilon}, \ \max(|F_J(\boldsymbol{\alpha})|, |F_K(\boldsymbol{\alpha})|) \leq P^{3/4+\varepsilon}\},$$

define $\mathfrak{f}(J)$ and $\mathfrak{f}(K)$ likewise, and put

$$\mathfrak{f} = \mathfrak{f}(I) \cup \mathfrak{f}(J) \cup \mathfrak{f}(K).$$

Similarly, let

$$\mathfrak{g}(I) = \{\boldsymbol{\alpha} \in \mathfrak{m} : |F_I(\boldsymbol{\alpha})| \leq P^{3/4+\varepsilon}, \ \min(|F_J(\boldsymbol{\alpha})|, |F_K(\boldsymbol{\alpha})|) > P^{3/4+\varepsilon}\},$$

define $\mathfrak{g}(J)$ and $\mathfrak{g}(K)$ likewise, and put

$$\mathfrak{g} = \mathfrak{g}(I) \cup \mathfrak{g}(J) \cup \mathfrak{g}(K).$$

Finally, define

$$\mathfrak{h} = \{\boldsymbol{\alpha} \in \mathfrak{m} : |F_i(\boldsymbol{\alpha})| > P^{3/4+\varepsilon} \text{ for } i = I, J, K\}.$$

The set $\mathfrak{e}$ can be handled quite easily. Using (2.15)–(2.18) and Lemma 2.4.1, we obtain

$$\begin{aligned}
\iint\limits_{\mathfrak{e}} |\mathcal{F}\mathcal{G}\mathcal{H}\mathcal{K}|\, d\boldsymbol{\alpha} &\ll Q\left(P^{3/4+\varepsilon}\right)^3 \iint\limits_{\mathcal{U}_{\mathfrak{e}}} \left(f^{10} + f^6H^4 + g^6H^4 + f^4g^6\right) d\boldsymbol{\alpha} \\
&\ll P^{s-13+\sigma_1+\sigma_2+9/4+3\varepsilon}\left(P^{17/3+\varepsilon} + P^{21/4+\varepsilon}\right) \\
&= o(P^{s-5}),
\end{aligned}$$

provided that $\sigma_1 + \sigma_2 < 1/12$, since ε can be chosen arbitrarily small.

The rational approximations provided by Lemma 2.4.2 allow us to incorporate major arc techniques along the lines of Brüdern [11] and [12] in dealing with the sets $\mathfrak{f}$, $\mathfrak{g}$, and $\mathfrak{h}$. For this we require some additional definitions and lemmata. Define

$$\mathcal{M}(q,a,b) = \{\boldsymbol{\alpha} \in [0,1]^2 : |q\alpha - a| < P^{-9/4} \text{ and } |q\beta - b| < P^{-5/4}\},$$

$$\mathcal{M} = \bigcup_{\substack{0\le a,b\le q<P^{3/4} \\ (q,a,b)=1}} \mathcal{M}(q,a,b),$$

$$S(q,a,b) = \sum_{x=1}^{q} e\left(\frac{ax^3+bx^2}{q}\right),$$

and

$$S_t^*(q) = \sum_{\substack{1\le a,b\le q \\ (q,a,b)=1}} \left|q^{-1}S(q,a,b)\right|^t.$$

Lemma 2.4.3. *For $t > 6$, we have*

$$\sum_{q\le X} S_t^*(q) \ll 1.$$

Proof. Using Lemma 10.4 of Wooley [61] and proceeding as in Lemma 2.11 of Vaughan [55], one sees that $S_t^*(q)$ is multiplicative, so

$$\sum_{q\le X} S_t^*(q) \le \prod_p \left(1 + \sum_{h=1}^{\infty} S_t^*(p^h)\right). \tag{2.19}$$

Whenever $(p^h,a,b) = 1$, we have

$$S(p^h,a,b) \ll p^{2h/3+\varepsilon}$$

by Theorem 7.1 of Vaughan [55], but in the case that $(b,p)=1$ it follows from Theorem 1 of Loxton and Vaughan [38] that in fact

$$S(p^h,a,b) \ll p^{h/2}.$$

Thus we have

$$\begin{aligned} S_t^*(p^h) &= p^{-ht} \sum_{\substack{1\le a,b\le p^h\\ (p,b)=1}} \left|S(p^h,a,b)\right|^t + p^{-ht} \sum_{\substack{1\le a,b\le p^h\\ (p^h,a,b)=1\\ (p,b)>1}} \left|S(p^h,a,b)\right|^t \\ &\ll p^{-ht}\big(p^{2h+ht/2} + p^{2h-1+2ht/3+t\varepsilon}\big), \end{aligned}$$

whence for $t>6$ we have

$$\sum_{h=1}^{\infty} S_t^*(p^h) \ll p^{-1-\delta}$$

for some $\delta>0$, and the result now follows immediately from (2.19). $\square$

Write

$$F(\boldsymbol{\alpha}) = \sum_{1\le x\le P} e(\alpha x^3 + \beta x^2) \tag{2.20}$$

and

$$v(\boldsymbol{\alpha}) = \int_0^P e(\alpha\gamma^3 + \beta\gamma^2)\, d\gamma. \tag{2.21}$$

The following lemma provides a useful refinement of [61], Lemma 9.2.

Lemma 2.4.4. *For $t>6$, we have*

$$\iint_{\mathcal{M}} |F(\boldsymbol{\alpha})|^t \, d\boldsymbol{\alpha} \ll P^{t-5}.$$

Proof. When $\boldsymbol{\alpha}\in\mathcal{M}(q,a,b)$, write $\boldsymbol{\xi}=(\xi_3,\xi_2)=(\alpha-a/q,\beta-b/q)$ and

$$V(\boldsymbol{\alpha}) = V(\boldsymbol{\alpha};q,a,b) = q^{-1}S(q,a,b)v(\boldsymbol{\xi}).$$

Then for $\boldsymbol{\alpha}\in\mathcal{M}(q,a,b)$ we have by Lemma 4.4 of Baker [5] that

$$F(\boldsymbol{\alpha}) = V(\boldsymbol{\alpha}) + O(q^{2/3+\varepsilon}).$$

Hence if $\mathcal{M}_1$ denotes the subset of $\mathcal{M}$ on which $|V(\boldsymbol{\alpha})| \leq q^{2/3+\varepsilon}$, then we have

$$\iint_{\mathcal{M}_1} |F(\boldsymbol{\alpha})|^t \, d\boldsymbol{\alpha} \ll \sum_{q \leq P^{3/4}} (q^{2/3+\varepsilon})^t P^{-7/2} \ll P^{t-5},$$

provided that $t > 9/2$. For $\boldsymbol{\alpha} \in \mathcal{M}_2 = \mathcal{M} \setminus \mathcal{M}_1$, we have $|V(\boldsymbol{\alpha})| > q^{2/3+\varepsilon}$ and hence $|F(\boldsymbol{\alpha})| \ll |V(\boldsymbol{\alpha})|$. Moreover, by Theorem 7.3 of Vaughan [55], we have

$$v(\boldsymbol{\xi}) \ll P(1 + P^2|\xi_2| + P^3|\xi_3|)^{-1/3} \ll P(1 + P^2|\xi_2|)^{-1/6}(1 + P^3|\xi_3|)^{-1/6},$$

and on combining this with Lemma 2.4.3 we obtain

$$\iint_{\mathcal{M}} |V(\boldsymbol{\alpha})|^t d\boldsymbol{\alpha} \ll P^{t-5} \sum_{q \leq P^{3/4}} S_t^*(q) \ll P^{t-5}$$

whenever $t > 6$. Thus we have

$$\iint_{\mathcal{M}_2} |F(\boldsymbol{\alpha})|^t d\boldsymbol{\alpha} \ll P^{t-5}$$

for $t > 6$, and this completes the proof. □

The sets $\mathfrak{f}$ and $\mathfrak{g}$ can now be handled with little difficulty by applying major arc treatments to one or two of the variables. The key observation is that Baker's Theorem (Lemma 2.4.2) allows us to bound an integral of $|F_i(\boldsymbol{\alpha})|^t$ over $\mathfrak{f}(i)^*$ or $\mathfrak{g}(j)^*$ $(j \neq i)$ in terms of the integral considered in the previous lemma.

Using (2.15)–(2.18) as on $\mathfrak{e}$, we obtain for some $i = I, J$, or K that

$$\iint_{\mathfrak{f}} |\mathcal{FGHK}| \, d\boldsymbol{\alpha} \ll Q \left(P^{3/4+\varepsilon}\right)^2 \iint_{\mathfrak{f}(i)^*} |F_i| \left(f^{10} + f^6 H^4 + g^6 H^4 + f^4 g^6\right) d\boldsymbol{\alpha}.$$

Then by Hölder's inequality we have

$$\iint_{\mathfrak{f}(i)^*} |F_i| \, f^{10} d\boldsymbol{\alpha} \ll \left(\iint_{\mathfrak{f}(i)^*} |F_i|^7 \, d\boldsymbol{\alpha} \right)^{1/7} \left(\iint_{\mathcal{U}_{\mathfrak{f}}} f^{10} d\boldsymbol{\alpha} \right)^{1/2} \left(\iint_{\mathcal{U}_{\mathfrak{f}}} f^{14} d\boldsymbol{\alpha} \right)^{5/14},$$

and by Lemma 2.4.1 we have

$$\iint_{\mathfrak{f}(i)^*} |F_i| \left(f^6 H^4 + g^6 H^4 + f^4 g^6\right) d\boldsymbol{\alpha} \ll P^{25/4+\varepsilon}.$$

Hence on using Lemmata 2.4.1, 2.4.2, and 2.4.4, together with a change of variables, we find that

$$\iint_{\mathfrak{f}} |\mathcal{FGHK}| \, d\boldsymbol{\alpha} \;\ll\; P^{s-13+\sigma_1+\sigma_2+3/2+2\varepsilon} \left(P^{19/3+\varepsilon} + P^{25/4+\varepsilon}\right) \;=\; o(P^{s-5}),$$

provided that $\sigma_1 + \sigma_2 < 1/6$.

Proceeding similarly but instead taking $u = 40/7$ in (2.17), we have for some $i \neq j$ among I, J, and K that

$$\begin{aligned}\iint\limits_{\mathfrak{g}} |\mathcal{FGHK}|\,d\boldsymbol{\alpha} &\ll QP^{3/4+\varepsilon} \iint\limits_{\mathfrak{g}(i)^*} |F_j|^2 \left(f^{10} + f^{\frac{40}{7}} H^{\frac{30}{7}} + g^{\frac{40}{7}} H^{\frac{30}{7}} + f^{\frac{30}{7}} g^{\frac{40}{7}}\right) d\boldsymbol{\alpha} \\ &\ll QP^{3/4+\varepsilon} \left(\iint\limits_{\mathfrak{g}(i)^*} |F_j|^7\,d\boldsymbol{\alpha} \right)^{2/7} \left(\mathcal{I}_1^{5/7} + \mathcal{I}_2^{5/7} + \mathcal{I}_3^{5/7} + \mathcal{I}_4^{5/7}\right),\end{aligned}$$

where

$$\mathcal{I}_1 = \iint\limits_{\mathcal{U}_{\mathfrak{g}}} f^{14} d\boldsymbol{\alpha}, \qquad \mathcal{I}_2 = \iint\limits_{\mathcal{U}_{\mathfrak{g}}} f^8 H^6 d\boldsymbol{\alpha},$$
$$\mathcal{I}_3 = \iint\limits_{\mathcal{U}_{\mathfrak{g}}} g^8 H^6 d\boldsymbol{\alpha}, \qquad \mathcal{I}_4 = \iint\limits_{\mathcal{U}_{\mathfrak{g}}} f^6 g^8 d\boldsymbol{\alpha}.$$

Thus we have

$$\iint\limits_{\mathfrak{g}} |\mathcal{FGHK}|\,d\boldsymbol{\alpha} \ll P^{s-13+\sigma_1+\sigma_2+3/4+\varepsilon} \left(P^7\right) = o(P^{s-5}),$$

provided that $\sigma_1 + \sigma_2 < 1/4$.

The set $\mathfrak{h}$ is somewhat more difficult to deal with, and it is here that we make use of the hypothesis that λ_I/λ_J and μ_I/μ_J are algebraic irrationals. We divide $\mathfrak{h}$ into two main components,

$$\mathfrak{h}_1 = \{\boldsymbol{\alpha} \in \mathfrak{h} : |\alpha| \geq P^{-9/4+\varepsilon}\} \quad \text{and} \quad \mathfrak{h}_2 = \mathfrak{h} \setminus \mathfrak{h}_1,$$

and we further subdivide $\mathfrak{h}_1^*$ and $\mathfrak{h}_2^*$ into $O\left((\log P)^2\right)$ dyadic subsets of the form

$$\mathfrak{h}_i(A,B) = \{\boldsymbol{\alpha} \in \mathfrak{h}_i^* : A < |F_I(\boldsymbol{\alpha})| \leq 2A,\ B < |F_J(\boldsymbol{\alpha})| \leq 2B\}.$$

We also write

$$\mathfrak{h}(A,B) = \mathfrak{h}_1(A,B) \cup \mathfrak{h}_2(A,B).$$

We now use a method introduced by Baker [4] to give an upper bound for the Lebesgue measure of $\mathfrak{h}_i(A,B)$. If $\boldsymbol{\alpha} \in \mathfrak{h}(A,B)$, then by Lemma 2.4.2 there exist natural numbers

$$q_I < P^{3+\varepsilon} A^{-3}, \quad q_J < P^{3+\varepsilon} B^{-3}, \quad q_K < P^{3/4} \tag{2.22}$$

and integers a_i, b_i with $(q_i, a_i, b_i) = 1$ for $i = I, J, K$ such that

$$|\lambda_I \alpha q_I - a_I| < P^{\varepsilon} A^{-3}, \quad |\mu_I \beta q_I - b_I| < P^{1+\varepsilon} A^{-3}; \tag{2.23}$$

$$|\lambda_J \alpha q_J - a_J| < P^{\varepsilon} B^{-3}, \quad |\mu_J \beta q_J - b_J| < P^{1+\varepsilon} B^{-3}; \tag{2.24}$$

and

$$|\lambda_K \alpha q_K - a_K| < P^{-9/4}, \quad |\mu_K \beta q_K - b_K| < P^{-5/4}. \tag{2.25}$$

Notice that the inequalities (2.23) and (2.24) restrict $\boldsymbol{\alpha}$ to lie in a box $\mathcal{B}_I$ about the point $(a_I/(\lambda_I q_I), b_I/(\mu_I q_I))$ with

$$\operatorname{meas}(\mathcal{B}_I) \ll q_I^{-2} P^{1+2\varepsilon} A^{-6} \tag{2.26}$$

and at the same time in a box $\mathcal{B}_J$ about $(a_J/(\lambda_J q_J), b_J/(\mu_J q_J))$ with

$$\operatorname{meas}(\mathcal{B}_J) \ll q_J^{-2} P^{1+2\varepsilon} B^{-6}. \tag{2.27}$$

We first obtain a lower bound for $q_I q_J$. As in the proof of Lemma 11.1 of Vaughan [55], it follows from (2.23) and (2.24) that for $\boldsymbol{\alpha} \in \mathfrak{h}_1$ we have

$$\left| \frac{\lambda_I}{\lambda_J} - \frac{a_I q_J}{a_J q_I} \right| \ll P^{-9/4},$$

whereas by a well-known theorem of Roth [50] we have

$$\left| \frac{\lambda_I}{\lambda_J} - \frac{a_I q_J}{a_J q_I} \right| \gg \frac{1}{|a_J q_I|^{2+\varepsilon}},$$

so that $|a_J q_I| \gg P^{9/8-\varepsilon}$. Similarly, for $\boldsymbol{\alpha} \in \mathfrak{h}_2$ we have

$$\frac{1}{|b_J q_I|^{2+\varepsilon}} \ll \left| \frac{\mu_I}{\mu_J} - \frac{b_I q_J}{b_J q_I} \right| \ll P^{-5/4},$$

and hence $|b_J q_I| \gg P^{5/8-\varepsilon}$. Thus on using (2.23) and (2.24) and recalling the definitions (2.12)–(2.14) we obtain

$$q_I q_J \gg \begin{cases} P^{9/8-\sigma_1-2\varepsilon}, & \text{if } \boldsymbol{\alpha} \in \mathfrak{h}_1 \\ P^{5/8-\sigma_2-2\varepsilon}, & \text{if } \boldsymbol{\alpha} \in \mathfrak{h}_2. \end{cases} \tag{2.28}$$

Next we observe that when $\boldsymbol{\alpha} \in \mathfrak{h}_1(A, B)$ there are $O(P^{9+3\varepsilon} A^{-9})$ corresponding triples (q_I, a_I, b_I) satisfying (2.22) and (2.23). Alternatively, there are $O(P^{9+3\varepsilon} B^{-9})$

triples (q_J, a_J, b_J) satisfying (2.22) and (2.24). On combining this with (2.26), (2.27), and (2.28) we obtain

$$\operatorname{meas}(\mathfrak{h}_1(A,B)) \ll P^{71/8+\sigma_1+7\varepsilon}(AB)^{-15/2}. \tag{2.29}$$

When $\boldsymbol{\alpha} \in \mathfrak{h}_2(A,B)$ we necessarily have $a_I = a_J = 0$ for P sufficiently large, so proceeding as above gives

$$\operatorname{meas}(\mathfrak{h}_2(A,B)) \ll P^{51/8+\sigma_2+6\varepsilon}(AB)^{-6}. \tag{2.30}$$

On applying Hölder's inequality and Lemma 2.4.1 as before and writing $L = (\log P)^2$, we find that for some A and B

$$\begin{aligned}\iint\limits_{\mathfrak{h}_1} |\mathcal{FGHK}|\, d\boldsymbol{\alpha} &\ll QL \iint\limits_{\mathfrak{h}_1(A,B)} |F_I F_J F_K| \big(f^{10} + f^{\frac{40}{7}} H^{\frac{30}{7}} + g^{\frac{40}{7}} H^{\frac{30}{7}} + f^{\frac{30}{7}} g^{\frac{40}{7}}\big) d\boldsymbol{\alpha} \\ &\ll QP^{\varepsilon} \left(\iint\limits_{\mathfrak{h}_1^*} |F_K|^{\frac{105}{16}}\, d\boldsymbol{\alpha} \right)^{\frac{16}{105}} \left(\iint\limits_{\mathfrak{h}_1(A,B)} |F_I F_J|^{\frac{15}{2}}\, d\boldsymbol{\alpha} \right)^{\frac{2}{15}} \big(P^9\big)^{\frac{5}{7}}.\end{aligned}$$

Thus by (2.29) and Lemma 2.4.4 we have

$$\iint\limits_{\mathfrak{h}_1} |\mathcal{FGHK}|\, d\boldsymbol{\alpha} \ll P^{s-13+\frac{45}{7}+\frac{5}{21}+\frac{2}{15}(\frac{71}{8})+\frac{17}{15}\sigma_1+\sigma_2+2\varepsilon} = o(P^{s-5}),$$

provided that $\frac{17}{15}\sigma_1 + \sigma_2 < \frac{3}{20}$.

Since $\mathfrak{h}_2$ is a thin strip along the β-axis, we save a factor of P^{σ_1} in the analysis leading to (2.15), but the treatment is otherwise similar to the above. On writing $Q' = P^{s-13+\sigma_2}$, we have

$$\begin{aligned}\iint\limits_{\mathfrak{h}_2} |\mathcal{FGHK}|\, d\boldsymbol{\alpha} &\ll Q'L \iint\limits_{\mathfrak{h}_2(A,B)} |F_I F_J F_K| \big(f^{10} + f^{\frac{40}{7}} H^{\frac{30}{7}} + g^{\frac{40}{7}} H^{\frac{30}{7}} + f^{\frac{30}{7}} g^{\frac{40}{7}}\big) d\boldsymbol{\alpha} \\ &\ll P^{s-13+\sigma_2+\varepsilon} \left(\iint\limits_{\mathfrak{h}_2^*} |F_K|^{\frac{42}{5}}\, d\boldsymbol{\alpha} \right)^{\frac{5}{42}} \left(\iint\limits_{\mathfrak{h}_2(A,B)} |F_I F_J|^{6}\, d\boldsymbol{\alpha} \right)^{\frac{1}{6}} \big(P^9\big)^{\frac{5}{7}},\end{aligned}$$

whence by (2.30) we obtain

$$\iint\limits_{\mathfrak{h}_2} |\mathcal{FGHK}|\, d\boldsymbol{\alpha} \ll P^{s-13+\frac{45}{7}+\frac{17}{42}+\frac{1}{6}(\frac{51}{8})+\frac{7}{6}\sigma_2+2\varepsilon} = o(P^{s-5}),$$

provided that $\frac{7}{6}\sigma_2 < \frac{5}{48}$. It is easily seen that these last two inequalities are less restrictive than the one appearing in condition (d) of Theorem 2.1.

2.5 The Major Arc

As it stands, the major arc $\mathfrak{M}$ is too large to allow satisfactory approximation of the exponential sums $f_i(\boldsymbol{\alpha})$, so we must do some pruning. Specifically, let W be a parameter at our disposal, and let

$$\mathfrak{N} = \{\boldsymbol{\alpha} : |\alpha| \leq WP^{-3} \text{ and } |\beta| \leq WP^{-2}\}. \tag{2.31}$$

Then as in Lemma 9.2 of Wooley [61], we have for $t > 9$ that

$$\iint\limits_{\mathfrak{M}\setminus\mathfrak{N}} |F_i(\boldsymbol{\alpha})|^t \, d\boldsymbol{\alpha} \ll W^{-\sigma} P^{t-5}$$

for $i = I, J, K$ and some $\sigma > 0$. Thus by using (2.17) and Lemma 2.4.1 as in the treatment of $\mathfrak{g}$ and $\mathfrak{h}$ in the previous section, we have for some $i = I, J$, or K that

$$\begin{aligned} \iint\limits_{\mathfrak{M}\setminus\mathfrak{N}} |\mathcal{FGHK}| \, d\boldsymbol{\alpha} &\ll P^{s-13} \left(\iint\limits_{\mathfrak{M}\setminus\mathfrak{N}} |F_i(\boldsymbol{\alpha})|^{21/2} \, d\boldsymbol{\alpha} \right)^{2/7} P^{45/7} \\ &\ll P^{s-5} W^{-\sigma'}. \end{aligned}$$

It remains to deal with the pruned major arc $\mathfrak{N}$. Let

$$v_i(\boldsymbol{\alpha}) = \int_0^P e(\lambda_i \alpha \gamma^3 + \mu_i \beta \gamma^2) \, d\gamma \tag{2.32}$$

and

$$w_i(\boldsymbol{\alpha}) = \int_R^P \rho\left(\frac{\log \gamma}{\log R}\right) e(\lambda_i \alpha \gamma^3 + \mu_i \beta \gamma^2) \, d\gamma, \tag{2.33}$$

where $\rho(x)$ is Dickman's function (see Vaughan [55], chapter 12). Then for $\boldsymbol{\alpha} \in \mathfrak{N}$, we obtain from Theorem 7.2 of [55] that

$$F_i(\boldsymbol{\alpha}) = v_i(\boldsymbol{\alpha}) + O(W)$$

and from Lemma 8.5 of [61] that

$$f_i(\boldsymbol{\alpha}) = w_i(\boldsymbol{\alpha}) + O(WP/\log P).$$

Now on taking $W = (\log P)^{1/4}$ it follows that

$$\iint\limits_{\mathfrak{N}} \mathcal{FGHK} \, d\boldsymbol{\alpha} = \iint\limits_{\mathfrak{N}} \left(\prod_{i=1}^{m+h-3} w_i(\boldsymbol{\alpha}) \right) \left(\prod_{i=m+h-2}^{s} v_i(\boldsymbol{\alpha}) \right) \mathcal{K}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} + O(P^{s-5} W^{-1}).$$

Furthermore, we may extend the integration over all of $\mathbb{R}^2$, as the bounds for v_i and w_i contained in Lemma 8.6 of [61] imply that

$$\iint\limits_{\mathbb{R}^2\setminus\mathfrak{N}} \left(\prod_{i=1}^{m+h-3} w_i(\boldsymbol{\alpha})\right)\left(\prod_{i=m+h-2}^{s} v_i(\boldsymbol{\alpha})\right)\mathcal{K}(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} \ll P^{s-5}W^{-1}.$$

Thus it remains to show that the singular integral

$$J = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\prod_{i=1}^{m+h-3} w_i(\boldsymbol{\alpha})\right)\left(\prod_{i=m+h-2}^{s} v_i(\boldsymbol{\alpha})\right)\mathcal{K}(\boldsymbol{\alpha})\,d\boldsymbol{\alpha}$$

satisfies $J \gg P^{s-5}$. Multiplying out, we have

$$J = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{\mathcal{B}^*} T^*(\boldsymbol{\gamma})\,e(F(\boldsymbol{\gamma})\alpha + G(\boldsymbol{\gamma})\beta)\,K(\alpha P^{-\sigma_1})K(\beta P^{-\sigma_2})\,d\boldsymbol{\gamma}\,d\alpha\,d\beta,$$

where

$$\mathcal{B}^* = [R,P]^{m+h-3}\times[0,P]^{n+3}$$

and

$$T^*(\boldsymbol{\gamma}) = \prod_{i=1}^{m+h-3}\rho\left(\frac{\log\gamma_i}{\log R}\right).$$

On making the change of variables

$$\boldsymbol{\gamma}' = \boldsymbol{\gamma}P^{-1},\quad \alpha' = \alpha P^{-\sigma_1},\quad \beta' = \beta P^{-\sigma_2}$$

and applying Fubini's Theorem, we obtain

$$J = P^{s+\sigma_1+\sigma_2}\int_{\mathcal{B}} T(\boldsymbol{\gamma})\,\hat{K}(F(\boldsymbol{\gamma})P^{3+\sigma_1})\hat{K}(G(\boldsymbol{\gamma})P^{2+\sigma_2})\,d\boldsymbol{\gamma}, \tag{2.34}$$

where we have written

$$\mathcal{B} = P^{-1}\mathcal{B}^*,\qquad T(\boldsymbol{\gamma}) = T^*(P\boldsymbol{\gamma}),$$

and

$$\hat{K}(t) = \int_{-\infty}^{\infty} e(\alpha t)K(\alpha)\,d\alpha.$$

Now by condition (c) of Theorem 2.1 and the argument of Lemma 6.2 of Wooley [61], we can find a non-singular solution $\boldsymbol{\eta}$ to the equations $F = G = 0$ such that each η_i is non-zero. Then, after replacing λ_i by $-\lambda_i$ if necessary and using homogeneity, we may assume that $\boldsymbol{\eta} \in (0,1)^s$ and hence that $\boldsymbol{\eta}$ lies in the interior of $\mathcal{B}$ when P is sufficiently large. Suppose that $6\eta_j\eta_k(\lambda_j\mu_k\eta_j - \lambda_k\mu_j\eta_k) \neq 0$, and consider the map $\phi : \mathbb{R}^s \to \mathbb{R}^s$ defined by

$$\phi_j = F(\boldsymbol{\gamma}), \quad \phi_k = G(\boldsymbol{\gamma}), \quad \text{and} \quad \phi_i = \gamma_i \quad (i \neq j, k). \tag{2.35}$$

By the inverse function theorem, there exist neighborhoods U of $\boldsymbol{\eta}$ and V of $\phi(\boldsymbol{\eta})$ such that ϕ maps U injectively onto V, and we may assume that $U \subset \mathcal{B}$. Now by (2.10) and the nonnegativity of ρ, the integrand in (2.34) is nonnegative, so we may restrict the integration over $\boldsymbol{\gamma}$ to the set U. Then on writing $\mathbf{z} = \phi(\boldsymbol{\gamma})$, where ϕ is as in (2.35), we have by the change of variables theorem that

$$J \geq P^{s+\sigma_1+\sigma_2} \int_V T(\phi^{-1}(\mathbf{z}))\hat{K}(z_j P^{3+\sigma_1})\hat{K}(z_k P^{2+\sigma_2}) \left| \frac{d\boldsymbol{\gamma}}{d\mathbf{z}} \right| d\mathbf{z}. \tag{2.36}$$

Since $\mathrm{meas}(V) \gg 1$, the projection of V onto z_j contains the interval $[0, \frac{1}{3}P^{-3-\sigma_1}]$, and the projection of V onto z_k contains the interval $[0, \frac{1}{3}P^{-2-\sigma_2}]$, provided that P is sufficiently large. Hence on restricting the range of integration in (2.36) and using (2.10) again, we obtain

$$J \gg P^{s+\sigma_1+\sigma_2} \int_{\mathcal{S}} T(\phi^{-1}(\mathbf{z}))d\mathbf{z},$$

where $\mathrm{meas}(\mathcal{S}) \gg P^{-5-\sigma_1-\sigma_2}$. Finally, on noting that $T(\boldsymbol{\gamma}) \gg \rho(1/\eta)^{m+h-3} \gg 1$ for $\boldsymbol{\gamma} \in \mathcal{B}$, we obtain $J \gg P^{s-5}$ as required. This completes the proof of Theorem 2.1.

2.6 A Discussion of Possible Improvements

Here we discuss the possibility of weakening some of the conditions imposed on the forms F and G in Theorems 2.1 and 2.2. In view of the discussion of Wooley [61], §5, where it is shown that many conditions similar to ours are essentially best possible for the corresponding problem on equations, our observations will leave something to be desired. Nevertheless, we can show that at least some minimal conditions are necessary to ensure the solubility of (2.1).

For example, let

$$F(\mathbf{x}) = \lambda^3 x_1^3 - x_2^3 \quad \text{and} \quad G(\mathbf{x}) = \mu^2 x_3^2 - x_4^2,$$

where λ and μ are positive real algebraic of degree 3 and 2, respectively, such that λ^3 and μ^2 are irrational. For instance, we may take $\lambda = 1 + \sqrt[3]{2}$ and $\mu = 1 + \sqrt{2}$. Then it follows easily from Liouville's Theorem that, for sufficiently small $\tau > 0$, neither of the inequalities

$$|F(\mathbf{x})| < \tau, \quad |G(\mathbf{x})| < \tau$$

has a non-trivial solution in rational integers. Of course, this example is easily generalized to produce forms $F_1, \ldots, F_t$ of degrees $k_1, \ldots, k_t$ in $2t$ variables that do not take arbitrarily small values. Therefore, we must minimally require either $s \geq 5$ total variables or at least 3 variables explicit in one of the two forms.

More realistically, in light of [72], Theorem 1, one might hope to be able to prove Theorem 2.1 with $s = 13$ but conditions (a) and (b) weakened so that F and G need only have 7 and 5 variables explicit, respectively, rather than 9 and 8. The latter numbers arise from the inequalities (2.6), on which the analytic argument in Sections 2.3–2.5 depends, but one may attempt to reduce these in the manner of [61] and [72] by using Theorem 2.2. Unfortunately, there are some difficulties with this approach in our situation. If F has exactly 7 or 8 variables explicit, then we may apply Theorem 2.2 to solve (2.1), but we must settle for the inferior values of σ_1 and σ_2 allowed by condition (d)(i) of that theorem, and we forfeit our estimate for the density of solutions. Moreover, if G has exactly 7 variables explicit and F has at least 10 variables explicit, then neither Theorem 2.1 nor Theorem 2.2 applies with $s = 13$. To avoid this difficulty, we may hope to reduce the number of zero coefficients required by condition (d)(ii) of the latter from 7 to 6, and we saw in Section 2.2 that a conditional result of this type could be obtained using hypothetical results on small solutions of cubic inequalities in 7 variables.

As mentioned in Section 2.1, condition (b) of Theorem 2.2 can be eliminated from the stated version of the theorem, but some form of it is likely to be necessary for any desirable refinement of (d)(ii). If a quantitative version of the result of Margulis [39] on the Oppenheim conjecture were available, then we could reduce the 5 to 3 in condition (b) of our hypothetical version of Theorem 2.2, provided we assumed additionally that G is not a multiple of a form with integer coefficients. However, the methods of [39] do not seem to hold much promise for obtaining such a result.

We can also investigate the possibility of reducing the total number of variables required. Although Theorem 2.1 could conceivably hold with as few as 5 variables, it does not seem possible for an analytic argument of the flavor given in Sections 2.3–2.5 to be successful with fewer than 11 variables. In the "ideal" situation that the first four mean values in Lemma 2.4.1 were bounded by $P^{5+\varepsilon}$, a simplified version of our analysis would allow us to prove a version of the theorem for $s \geq 12$, possibly with a slightly different range of permissible values for σ_1 and σ_2.

Next we note that the existence of a non-trivial real solution to the equations $F = G = 0$ is a necessary condition for the system (2.1) to have infinitely many integer solutions. For, if the latter holds, then for arbitrary $\tau > 0$ we can obtain (by rescaling an integer solution $\mathbf{x}$ with $\max |x_i|$ sufficiently large) a real solution

$\boldsymbol{\eta}(\tau) \in [-1,1]^s$ of the inequalities $|F| < \tau$, $|G| < \tau$ such that $|\eta_i| = 1$ for some i. But the set

$$\mathcal{S} = \{\boldsymbol{\eta} \in [-1,1]^s : |\eta_i| = 1 \text{ for some } i\}$$

is compact, whence its image in $\mathbb{R}^2$ under the continuous map ϕ defined by F and G is compact. Hence $\phi(\mathcal{S})$ must contain the limit point $(0,0)$, which shows that the equations $F = G = 0$ have a non-trivial real solution.

Now let p be a prime with $p \equiv 1 \pmod 3$, let c be a cubic nonresidue $(\mathrm{mod}\ p)$, and consider the forms

$$\begin{aligned} F(\mathbf{x}) &= \sqrt{2}x_1^3 + x_2^3 + \cdots + x_7^3 + (x_8^3 + cx_9^3) + p(x_{10}^3 + cx_{11}^3) + p^2(x_{12}^3 + cx_{13}^3), \\ G(\mathbf{x}) &= \sqrt{2}x_1^2 + x_2^2 + \cdots + x_7^2 + x_8^2. \end{aligned}$$

It is easily checked that F and G satisfy all the conditions of Theorem 2.1, except that all real solutions to the simultaneous equations $F = G = 0$ are singular. Moreover, the discussion of example (5.1) in Wooley [61] shows that the simultaneous inequalities

$$|F(\mathbf{x})| < 1, \quad |G(\mathbf{x})| < 1$$

have no nontrivial integer solutions. Therefore, condition (c) of Theorem 2.1 cannot be weakened.

We conclude with some remarks on the assumption regarding algebraic irrational coefficient ratios in Theorem 2.1. First of all, if neither F nor G is a multiple of a form with integer coefficients and all the coefficients of F and G are nonzero, then it is easy to see that there is a pair of indices i and j such that both λ_i/λ_j and μ_i/μ_j are irrational. Next, if exactly one of the forms is a multiple of an integral form and this form has no zero coefficients, then we can solve the problem by obtaining a lower bound for the integral

$$R_1(P) = \int_{-\infty}^{\infty} \int_0^1 \mathcal{F}(\boldsymbol{\alpha})\mathcal{G}(\boldsymbol{\alpha})\mathcal{H}(\boldsymbol{\alpha})K(\alpha P^{-\sigma_1})\, d\beta\, d\alpha$$

or

$$R_2(P) = \int_{-\infty}^{\infty} \int_0^1 \mathcal{F}(\boldsymbol{\alpha})\mathcal{G}(\boldsymbol{\alpha})\mathcal{H}(\boldsymbol{\alpha})K(\beta P^{-\sigma_2})\, d\alpha\, d\beta,$$

as the case may be, using a simplified version of our analysis, along with techniques from the one-dimensional Hardy-Littlewood and Davenport-Heilbronn methods. If F and G are both multiples of integral forms, then we may simply apply the argument of Wooley [72] to deduce Theorem 2.1. Thus in particular we observe that if all the coefficients of F and G are algebraic and nonzero, then no irrationality assumption on the coefficients is needed.

The algebraicity assumption allows us to use Roth's Theorem in Section 2.4 to obtain the lower bounds (2.28), which are critical to our analysis of the sets $\mathfrak{h}_i(A, B)$. The preferred approach to (2.28) would involve restricting P in terms of the denominators of simultaneous rational approximations $\lambda_I/\lambda_J \sim a/q$ and $\mu_I/\mu_J \sim b/q$ and then combining these approximations with (2.23) and (2.24), in analogy with the proof of [55], Lemma 11.1. However, a difficulty arises from the possibility that (a, q) or (b, q) may be large, even though we can ensure that $(q, a, b) = 1$. It transpires that in this problematic case we can reduce the task to one of obtaining small solutions to "mixed" systems of the form

$$|F(\mathbf{x})| < (\max |x_i|)^{-\sigma_1}, \quad \sum_{i=1}^{s} b_i x_i^2 = 0$$

or

$$|G(\mathbf{x})| < (\max |x_i|)^{-\sigma_2}, \quad \sum_{i=1}^{s} a_i x_i^3 = 0,$$

where the a_i and b_i are integers. Under suitable conditions, the number of solutions to these systems can be estimated as described above, using integrals like $R_1(P)$ and $R_2(P)$. However, in order to obtain bounds for the solutions in terms of the coefficients of the forms, we must now keep track of constants that were previously left implicit, and this would seem to require additional information regarding the nature of a real solution to the corresponding system of equations.

CHAPTER III

Mean Values of Multiple Exponential Sums

3.1 Overview

In this chapter, we obtain estimates for mean values of certain multiple exponential sums over smooth numbers by extending the ideas of Vaughan [53] and Wooley [61], [69]. When P and R are positive integers, write

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

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

$$f(\boldsymbol{\alpha};P,R) = \sum_{x,y\in\mathcal{A}(P,R)} e(\alpha_0 x^k + \alpha_1 x^{k-1} y + \cdots + \alpha_k y^k). \tag{3.1}$$

Further, define the mean value

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

which, by orthogonality, counts the number of solutions of the auxiliary system

$$\sum_{m=1}^{s} (x_m^{k-i} y_m^i - \tilde{x}_m^{k-i} \tilde{y}_m^i) = 0 \quad (0 \leq i \leq k) \tag{3.2}$$

with

$$x_m, y_m, \tilde{x}_m, \tilde{y}_m \in \mathcal{A}(P,R) \quad (1 \leq m \leq s). \tag{3.3}$$

When R is a power of P, one has the lower bound

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

so when $s \geq \frac{1}{2}k(k+1)$ we hope to obtain upper bounds that are not too much larger than $P^{4s-k(k+1)}$. If for every $\varepsilon > 0$ there exists $\eta = \eta(s,k,\varepsilon)$ such that the estimate

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

holds whenever $R \le P^{\eta}$, then we call Δ_s an admissible exponent.

After discussing some preliminary results in Section 3.2, we develop our version of the Vaughan-Wooley iterative method in Sections 3.3 and 3.4. In Section 3.5, we are able to establish the following result using only single efficient differencing.

Theorem 3.1. *Let $k \ge 2$ be a positive integer, and put $r = \left[\frac{k+1}{2}\right]$. Further, write*

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

and let s be a positive integer with $s \ge s_1$. Then the exponent

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

is admissible.

Notice in particular that if $s \ge 2k^2 \log k$ then admissible exponents obtained from Theorem 3.1 satisfy $\Delta_s \sim k^2 e^{-s/k^2} \le 1$.

Finally, in Section 3.6, we make full use of the repeated efficient differencing apparatus to obtain the following sharper result.

Theorem 3.2. *Write $r = \left[\frac{k+1}{2}\right]$, and put*

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

Further, define

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

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

Notice that the admissible exponents obtained from Theorem 3.2 decay in many cases of interest roughly like $k^2 e^{-3s/2k^2}$, whereas those obtained from Theorem 3.1 decay only like $k^2 e^{-s/k^2}$. Therefore, we will primarily rely on Theorem 3.2 when discussing the various applications of our mean value estimates for large k in Chapter 4. On the other hand, the ideas underlying Theorem 3.1 often suffice for smaller values of k, as we shall see in Chapter 5.

This chapter and the next are based on the author's submitted manuscript [46].

3.2 Preliminary Lemmata

Before embarking on the proofs of our mean value estimates, we need to make some preliminary observations. We start by showing that solutions of (3.2) in which some x_j and y_j or some $\tilde{x}_j$ and $\tilde{y}_j$ have a large common factor can effectively be ignored. When $\gamma > 0$, let $S_s(P,R;\gamma)$ be the number of solutions of (3.2) with $(x_j, y_j) \le P^\gamma$ and $(\tilde{x}_j, \tilde{y}_j) \le P^\gamma$ for all j.

Lemma 3.2.1. *For every $\gamma > 0$, one has $S_s(P,R) \ll P^{2s+\varepsilon} + S_s(P,R;\gamma)$.*

Proof. Write $S'_s(P,R;\gamma)$ for the number of solutions of (3.2) with $(x_j,y_j) > P^\gamma$ or $(\tilde{x}_j,\tilde{y}_j) > P^\gamma$ for some j, so that $S_s(P,R) = S_s(P,R;\gamma) + S'_s(P,R;\gamma)$. Then we have

$$S'_s(P,R;\gamma) = \sum_{d>P^\gamma} \int_{\mathbb{T}^{k+1}} f(d^k\boldsymbol{\alpha};P/d,R) f(-\boldsymbol{\alpha};P,R) |f(\boldsymbol{\alpha};P,R)|^{2s-2}\, d\boldsymbol{\alpha}. \tag{3.4}$$

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

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

If $\lambda_s \le 2s$, then we are done, so we may assume that $\lambda_s > 2s$. By applying Hölder's inequality to (3.4), we obtain

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

from which we deduce that

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

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

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

$$f_j(x_1,\dots,x_t) \equiv u_j \pmod{q^{k-j+1}p^{j-1}} \quad (1 \le j \le t) \tag{3.5}$$

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

$$J_t(\mathbf{f};\mathbf{x}) = \det\left(\frac{\partial f_i}{\partial x_j}(\mathbf{x}) \right)_{1\le i,j\le t}. \tag{3.6}$$

Lemma 3.2.2. *Suppose that $f_1, \dots, f_{2r} \in \mathbb{Z}[x_1, \dots, x_{2r}]$ have degrees bounded in terms of k. Then whenever $2r \le k+1$ we have*

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

Proof. Write $\tilde{q} = q/(q,p)$ and $\tilde{p} = p/(q,p)$, so that $(\tilde{q},\tilde{p}) = 1$. Then by considering the jth congruence in (3.5) modulo $\tilde{q}^{k-j+1}$, we obtain from Lemma 2.2 of Wooley [69] that the number of solutions modulo $\tilde{q}^k$ is $O_{\varepsilon,k}(\tilde{q}^{r(2r-1)+\varepsilon})$. Similarly, the number of solutions modulo $\tilde{p}^k$ is $O_{\varepsilon,k}(\tilde{p}^{r(2r-1)+\varepsilon})$. Hence by the Chinese Remainder Theorem the number of solutions modulo $\tilde{q}^k\tilde{p}^k$ is $O_{\varepsilon,k}((\tilde{p}\tilde{q})^{r(2r-1)+\varepsilon})$. Trivially, each of these solutions lifts in at most $(q,p)^{4kr}$ ways to $\mathbb{Z}/(q^kp^k)$, and the lemma follows immediately. □

We now develop some notation for analyzing real singular solutions of systems such as (3.2). Let $\psi_1, \dots, \psi_{2r}$ be non-trivial polynomials in $\mathbb{Z}[x,y]$ of total degree at most k. When $\mathcal{I}, \mathcal{J} \subset \{1,2,\dots,2r\}$ with $\operatorname{card}(\mathcal{J}) = 2\operatorname{card}(\mathcal{I})$ and $\mathbf{z}, \mathbf{w} \in \mathbb{Z}^{2r}$, define the Jacobian

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

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

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

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

Lemma 3.2.3. *Suppose that $\psi_1,\dots,\psi_{2r}$ satisfy the condition that $d_{1,2}$ is non-trivial and $\deg_w(d_{i,j}) < \deg_w(d_{i',j'})$ whenever $i+j < i'+j'$. Then we have*

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

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

$$d_{1,2}(z_i,w_i;\boldsymbol{\psi}) = 0 \tag{3.7}$$

for $i = 1,\dots,2r$. For a $4r$-tuple counted by $\mathcal{T}_0(P;\boldsymbol{\psi})$ and a given i, there are at most $O(P)$ choices for z_i and w_i satisfying (3.7), since we have assumed that $d_{1,2}$ is non-trivial, and it follows that $\operatorname{card}(\mathcal{T}_0(P;\boldsymbol{\psi})) \ll P^{2r}$.

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

$$J(\mathcal{I}, \mathcal{J}_{2d}; \boldsymbol{\psi}) \neq 0 \tag{3.8}$$

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

$$J(\mathcal{I} \cup \{i\}, \mathcal{J}_{2d+2}; \boldsymbol{\psi}) = 0 \tag{3.9}$$

for all $i \in \mathcal{J}_{2r} \setminus \mathcal{I}$. Consider a $4r$-tuple counted by $\mathcal{T}_d(P; \boldsymbol{\psi})$, where $1 \leq d \leq r-1$. There are $O(1)$ choices for $\mathcal{I}$ and $O(P^{2d})$ choices for the z_i and w_i with $i \in \mathcal{I}$. Now we fix $i \in \mathcal{J}_{2r} \setminus \mathcal{I}$ and expand the determinant in (3.9) using 2×2 blocks along the rows containing z_i and w_i. Then on using (3.8), together with our hypothesis on $\boldsymbol{\psi}$, we see that the relation (3.9) is a non-trivial polynomial equation in the variables z_i and w_i and hence has $O(P)$ solutions. Thus we have

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

and hence

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

as desired. □

Finally, we recall an estimate of Wooley [62] for the number of integers in an interval with a given square-free kernel s_0.

Lemma 3.2.4. *Suppose that L is a positive real number and that r is a positive integer with $\log r \ll \log L$. Then for each $\varepsilon > 0$, one has*

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

Proof. This is Lemma 2.1 of Wooley [62]. □

3.3 The Fundamental Lemma

For $0 \leq i \leq k$, let $\psi_i(z, w; \mathbf{c})$ be polynomials with integer coefficients in the variables $z, w, c_1, \ldots, c_u$ and satisfying the conditions of Lemma 3.2.3. Further, suppose that C_i and C_i' satisfy $1 \leq C_i' \leq C_i \ll P$, write $\tilde{C} = \prod_{i=1}^u C_i$, and let $D_i(\mathbf{c})$ be polynomials with total degrees bounded in terms of k such that $D_i(\mathbf{c}) \neq 0$ for $C_i' \leq c_i \leq C_i$. Throughout the remainder of this chapter, ε, η, and γ will denote small positive numbers, whose values may change from statement to statement. Generally, η and γ will be chosen sufficiently small in terms of ε, and the implicit constants in our analysis may depend at most on ε, η, γ, s, and k. Since our methods will involve only a finite number of steps, all implicit constants that arise remain under control, and the values assumed by η and γ throughout the arguments remain uniformly bounded away from zero.

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

$$\begin{aligned}\sum_{n=1}^{r} \eta_n(\psi_i(z_n,w_n;\mathbf{c}) - \psi_i(\tilde{z}_n,\tilde{w}_n;\mathbf{c}))&\\ + D_i(\mathbf{c})\sum_{m=1}^{s}(x_m^{k-i}y_m^i - \tilde{x}_m^{k-i}\tilde{y}_m^i) = 0 \quad (0 \le i \le k)&\end{aligned} \tag{3.10}$$

with

$$x_m, y_m, \tilde{x}_m, \tilde{y}_m \in \mathcal{A}(Q,R) \quad (1 \le m \le s), \tag{3.11}$$

$$(x_m,y_m) \le P^\gamma \quad \text{and} \quad (\tilde{x}_m,\tilde{y}_m) \le P^\gamma \quad (1 \le m \le s), \tag{3.12}$$

$$1 \le z_n, w_n, \tilde{z}_n, \tilde{w}_n \le P \quad \text{and} \quad \eta_n \in \{\pm 1\} \quad (1 \le n \le r), \tag{3.13}$$

and

$$C_j' \le c_j \le C_j \quad (1 \le j \le u). \tag{3.14}$$

Further, write $\tilde{S}_{s,r}(P,Q,R;\boldsymbol{\psi})$ for the number of solutions of (3.10) with (3.11), (3.12), (3.13), (3.14), and

$$J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c}) \ne 0 \quad \text{and} \quad J_{2r}(\tilde{\mathbf{z}},\tilde{\mathbf{w}};\mathbf{c}) \ne 0, \tag{3.15}$$

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

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

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

$$\begin{aligned}\sum_{n=1}^{r} \eta_n(\psi_i(z_n,w_n;\mathbf{c}) - \psi_i(\tilde{z}_n,\tilde{w}_n;\mathbf{c}))&\\ + D_i(\mathbf{c})q^{k-i}p^i\sum_{m=1}^{s}(u_m^{k-i}v_m^i - \tilde{u}_m^{k-i}\tilde{v}_m^i) = 0 \quad (0 \le i \le k)&\end{aligned} \tag{3.16}$$

with (3.13), (3.14),

$$P^\theta < p, q \le P^\theta R \quad \text{and} \quad (q,p) \le P^\gamma, \tag{3.17}$$

$$u_m, v_m, \tilde{u}_m, \tilde{v}_m \in \mathcal{A}(QP^{-\theta}, R) \quad (1 \leq m \leq s), \tag{3.18}$$

$$(u_m, v_m) \leq P^\gamma \quad \text{and} \quad (\tilde{u}_m, \tilde{v}_m) \leq P^\gamma \quad (1 \leq m \leq s), \tag{3.19}$$

and

$$(J_{2r}(\mathbf{z}, \mathbf{w}; \mathbf{c}), pq) = (J_{2r}(\tilde{\mathbf{z}}, \tilde{\mathbf{w}}; \mathbf{c}), pq) = 1. \tag{3.20}$$

Lemma 3.3.1. *Given $\varepsilon > 0$, there exists a positive number $\eta = \eta(\varepsilon, s, k)$ such that whenever $R \leq P^\eta$ one has*

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

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

(i) Suppose that $S_1 \geq S_2$, so that $S_{s,r}(P, Q, R; \boldsymbol{\psi}) \leq 2S_1$. By Lemma 3.2.3, we see that there are $O(P^{3r-1})$ permissible choices for $\mathbf{z}, \mathbf{w}, \tilde{\mathbf{z}}$, and $\tilde{\mathbf{w}}$. Now let

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

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

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

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

(ii) Suppose that $S_2 \geq S_1$, so that $S_{s,r}(P, Q, R; \boldsymbol{\psi}) \leq 2S_2$. By rearranging variables, we see that $S_{s,r}(P, Q, R; \boldsymbol{\psi}) \ll S_3$, where S_3 denotes the number of solutions of (3.10) with (3.11), (3.12), (3.13), and (3.14), and $J_{2r}(\mathbf{z}, \mathbf{w}; \mathbf{c}) \neq 0$. Then by using the Cauchy-Schwarz inequality as in the corresponding argument of Wooley [69] to manipulate the underlying mean values, we see that

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

where S_4 denotes the number of solutions of (3.10) with (3.11), (3.12), (3.13), (3.14), $J_{2r}(\mathbf{z}, \mathbf{w}; \mathbf{c}) \neq 0$, and $J_{2r}(\tilde{\mathbf{z}}, \tilde{\mathbf{w}}; \mathbf{c}) \neq 0$.

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

$$x_j\ \mathcal{D}(P^\theta)\ J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c}) \quad \text{or} \quad \tilde{x}_j\ \mathcal{D}(P^\theta)\ J_{2r}(\tilde{\mathbf{z}},\tilde{\mathbf{w}};\mathbf{c}) \tag{3.21}$$

or

$$y_j\ \mathcal{D}(P^\theta)\ J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c}) \quad \text{or} \quad \tilde{y}_j\ \mathcal{D}(P^\theta)\ J_{2r}(\tilde{\mathbf{z}},\tilde{\mathbf{w}};\mathbf{c}) \tag{3.22}$$

for some j, and let S_6 denote the number of solutions for which neither (3.21) nor (3.22) holds for any j. Then we have

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

and we divide into further cases.

(iii) Suppose that $S_5 \ge S_6$, and further suppose that (3.21) holds. Write

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

and let

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

where

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

Then

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

where

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

By using the Cauchy-Schwarz inequality and considering the underlying Diophantine equations as in [69], we deduce that

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

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

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

with (3.11), (3.12), (3.13), (3.14), and

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

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

Write

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

and

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

Then

$$S_{s,r}(P,Q,R;\boldsymbol{\psi})\ll\sum_{\mathbf{c},\boldsymbol{\eta}}\int_{\mathbb{T}^{k+1}}|\mathcal{G}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha})^2f_{\mathbf{c}}(\boldsymbol{\alpha};Q,R)^{2s-2}|\,d\boldsymbol{\alpha}.\tag{3.23}$$

By Cauchy's inequality, we have

$$|\mathcal{G}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha})|^2\le\mathcal{H}_{1,\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha})\mathcal{H}_{2,\mathbf{c}}(\boldsymbol{\alpha}),\tag{3.24}$$

where

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

and

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

Now by interchanging the order of summation and using Cauchy's inequality together with Lemma 3.2.4 as in [69], we obtain

$$\begin{aligned}
\mathcal{H}_{2,\mathbf{c}}(\boldsymbol{\alpha}) &= \sum_{g \le Q} \left| \sum_{\substack{x,y \le Q \\ s_0(x)=g}} \sum_{\substack{d \le P^\theta \\ d \le Q/x}} e\left(\sum_{i=0}^{k} \alpha_i D_i(\mathbf{c})(dx)^{k-i} y^i \right) \right|^2 \\
&\ll \sum_{g \le Q} Q^{1+\varepsilon} \sum_{\substack{x,y \le Q \\ s_0(x)=g}} P^\theta Q/x \\
&\ll Q^3 P^{\theta+\varepsilon}. \qquad (3.25)
\end{aligned}$$

Thus an application of Hölder's inequality in (3.23) gives

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

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

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

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

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

and the result follows in the case where (3.21) holds. The case where (3.22) holds is handled in exactly the same manner.

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

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

Then, since neither (3.21) nor (3.22) holds, we have $q > P^\theta$ and $p > P^\theta$. Thus we can find divisors q_j of x_j and p_j of y_j such that $P^\theta < q_j, p_j \le P^\theta R$ and $(q_j p_j, J_{2r}(\mathbf{z}, \mathbf{w}; \mathbf{c})) = 1$, and we proceed similarly with the $\tilde{x}_j$ and $\tilde{y}_j$, except that we replace $J_{2r}(\mathbf{z},\mathbf{w};\mathbf{c})$ by $J_{2r}(\tilde{\mathbf{z}},\tilde{\mathbf{w}};\mathbf{c})$. Hence we see that $S_6 \ll V_1$, where V_1 denotes the number of solutions of

$$\sum_{n=1}^{r}\eta_n(\psi_i(z_n,w_n;\mathbf{c})-\psi_i(\tilde{z}_n,\tilde{w}_n;\mathbf{c}))$$
$$+\ D_i(\mathbf{c})\sum_{j=1}^{s}((q_ju_j)^{k-i}(p_jv_j)^i-(\tilde{q}_j\tilde{u}_j)^{k-i}(\tilde{p}_j\tilde{v}_j)^i)=0 \quad (0\le i\le k).$$

with (3.13), (3.14), and for $1 \le j \le s$

$$P^\theta < q_j,p_j,\tilde{q}_j,\tilde{p}_j \le P^\theta R, \qquad (q_j,p_j),\ (\tilde{q}_j,\tilde{p}_j)\le P^\gamma, \tag{3.26}$$

$$u_j,v_j,\tilde{u}_j,\tilde{v}_j \in \mathcal{A}(QP^{-\theta},R), \qquad (u_j,v_j),\ (\tilde{u}_j,\tilde{v}_j)\le P^\gamma,$$

and

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

Now write

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

and

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

where

$$\mathbf{q}_j\mathbf{p}_j\boldsymbol{\alpha}=(\alpha_0q_j^k,\alpha_1q_j^{k-1}p_j,\ldots,\alpha_kp_j^k) \quad\text{and}\quad \tilde{\mathbf{q}}_j\tilde{\mathbf{p}}_j\boldsymbol{\alpha}=(\alpha_0\tilde{q}_j^k,\alpha_1\tilde{q}_j^{k-1}\tilde{p}_j,\ldots,\alpha_k\tilde{p}_j^k).$$

Then we have

$$V_1 \le \sum_{\mathbf{c},\boldsymbol{\eta}}\int_{\mathbb{T}^{k+1}}\sum_{\mathbf{q},\mathbf{p},\tilde{\mathbf{q}},\tilde{\mathbf{p}}} F_{\mathbf{c},\boldsymbol{\eta},\pi}(\boldsymbol{\alpha};P,R)F_{\mathbf{c},\boldsymbol{\eta},\tilde{\pi}}(-\boldsymbol{\alpha};P,R)\prod_{i=1}^{s}\mathcal{F}_{\mathbf{c},j}(\boldsymbol{\alpha})\,d\boldsymbol{\alpha}, \tag{3.27}$$

where

$$\pi=q_1\cdots q_sp_1\cdots p_s \quad\text{and}\quad \tilde{\pi}=\tilde{q}_1\cdots\tilde{q}_s\tilde{p}_1\cdots\tilde{p}_s,$$

and where the sum is over $\mathbf{q}$, $\mathbf{p}$, $\tilde{\mathbf{q}}$, $\tilde{\mathbf{p}}$ satisfying (3.26). Let

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

and let $Y_{\mathbf{c},\boldsymbol{\eta},j}(\boldsymbol{\alpha})$ be the analogous function for the $\tilde{q}_j$ and $\tilde{p}_j$. Then by (3.27) and two applications of Hölder's inequality (as in [69]), we obtain

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

Now we observe that

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

where $W(P,Q,R,q,p)$ denotes the number of solutions of (3.16) with (3.13), (3.14), (3.18), (3.19), and (3.20). Thus we have

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

whence by Hölder's inequality

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

and this completes the proof of the lemma. □

The following modification of Lemma 3.3.1 may be more useful for smaller values of k.

Lemma 3.3.2. *Given $\varepsilon > 0$, there exists a positive number $\eta = \eta(\varepsilon, s, k)$ such that whenever $R \le P^\eta$ one has*

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

Proof. The only change occurs in part (iii) of the proof, where the number of solutions counted by S_5 is estimated. Substituting the bounds (3.24) and (3.25) into (3.23), we obtain

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

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

Now let $\tilde{T}_{s,r}(P,Q,R,\theta;\boldsymbol{\psi})$ denote the number of solutions of (3.16) with (3.13), (3.14), (3.17), (3.18), (3.19) and also

$$z_n \equiv \tilde{z}_n \pmod{q^k p^k} \quad \text{and} \quad w_n \equiv \tilde{w}_n \pmod{q^k p^k} \quad (1 \le n \le r). \tag{3.28}$$

Lemma 3.3.3. *Given $\varepsilon > 0$, there exists a positive number $\gamma_0 = \gamma_0(\varepsilon, s, k)$ such that whenever $\gamma \le \gamma_0$ one has*

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

Proof. When q and p satisfy (3.17), let $\mathcal{B}_{q,p}(\mathbf{u};\mathbf{c},\boldsymbol{\eta})$ denote the set of solutions $(\mathbf{z},\mathbf{w})$ of the system of congruences

$$\Upsilon_i(\mathbf{z},\mathbf{w};\mathbf{c},\boldsymbol{\eta}) \equiv u_i \pmod{q^{k-i}p^i} \quad (0 \le i \le k) \tag{3.29}$$

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

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

By Lemma 3.2.2 we have

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

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

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

so for each i we can classify the solutions of (3.16) according to the common residue class modulo $q^{k-i}p^i$ of $\Upsilon_i(\mathbf{z},\mathbf{w};\mathbf{c},\boldsymbol{\eta})$ and $\Upsilon_i(\tilde{\mathbf{z}},\tilde{\mathbf{w}};\mathbf{c},\boldsymbol{\eta})$. Let

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

Then

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

where

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

and

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

Now by Cauchy's inequality,

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

and thus

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

This completes the proof. □

3.4 Efficient Differencing

Define the difference operator Δ_j^* recursively by

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

and

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

with the convention that

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

Further, write

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

where

$$h_i'=h_i(m_in_i)^k\quad\text{and}\quad g_i'=g_i(m_in_i)^k,\tag{3.30}$$

and put

$$r_j = \left[\frac{k-j+1}{2}\right]. \tag{3.31}$$

Our first task is to show that the polynomials $\psi_{i,j}$ satisfy the conditions of Lemma 3.2.3, so that the results of the previous section may be applied. We start by expressing Δ_j^* in terms of the more familiar difference operators Δ_j defined by

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

and

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

For simplicity, we introduce the functions

$$\chi_{i,j}(z,w;\mathbf{h};\mathbf{g}) = \Delta_j^*(z^{k-i}w^i;h_1,\ldots,h_j;g_1,\ldots,g_j) \tag{3.32}$$

and observe that

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

where $\mathbf{h}'$ and $\mathbf{g}'$ are defined by (3.30). As in Section 3.2, we write $\mathcal{J}_d$ for the set $\{1,\ldots,d\}$, and also write $\tilde{\mathcal{A}}_d$ for the set $\mathcal{J}_d \setminus \mathcal{A}$. When $\mathcal{A} = \{i_1,\ldots,i_m\} \subset \mathcal{J}_j$ with $i_1 < \cdots < i_m$, define

$$q_m^{(i)}(w;\mathbf{g},\mathcal{A}) = \Delta_m(w^i;g_{i_1},\ldots,g_{i_m}), \tag{3.33}$$

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

$$p_t^{(i)}(z;\mathbf{h},\mathcal{A},\mathcal{B}) = \Delta_t((z+h_{i_1}+\cdots+h_{i_m})^{k-i};h_{j_1},\ldots,h_{j_t}). \tag{3.34}$$

Lemma 3.4.1. *We have*

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

Proof. We fix i, $\mathbf{h}$, and $\mathbf{g}$ and proceed by induction on j. For brevity, we write $\chi_{i,j}(z,w)$, $q_m(w;\mathcal{A})$, and $p_t(z;\mathcal{A},\mathcal{B})$ for the functions defined by (3.32), (3.33), and (3.34), respectively. For $j=0$ we have

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

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

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

so by the inductive hypothesis we obtain

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

where

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

The above expression can be rewritten as

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

so we have

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

and the lemma follows. □

Now we show that the 2×2 Jacobians satisfy the condition imposed in Lemma 3.2.3.

Lemma 3.4.2. *Suppose that $0 \le j < k$ and $i_1 < i_2 \le k-j$ Then we have*

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

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

Proof. When $i < k-j$, we have by Lemma 3.4.1 that

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

and

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

and we recall (see for example Exercise 2.1 of Vaughan [55]) that

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

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

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

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

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

and the lemma follows in this case on noting that

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

Now if $i = k-j$ we obtain from Lemma 3.4.1 that

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

and

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

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

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

and this completes the proof. $\square$

We now consider the effect of substituting $\psi_{i,j}(z, w; \mathbf{h}, \mathbf{g}; \mathbf{m}, \mathbf{n})$ for $\psi_i(z, w; \mathbf{c})$ in the analysis of Section 3.3. For $1 \leq j \leq k$, suppose that $0 \leq \phi_j \leq 1/2k$, and put

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

Further, write

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

We replace (3.14) by the conditions

$$1 \leq h_i, g_i \leq H_i \quad (1 \leq i \leq j), \tag{3.35}$$

$$M_i < m_i, n_i \le M_i R, \quad \text{and} \quad (m_i, n_i) \le P^\gamma \quad (1 \le i \le j), \tag{3.36}$$

and take

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

On replacing h_i by $h_i(m_i n_i)^k$ and g_i by $g_i(m_i n_i)^k$ in the above results, we see that $\psi_{0,j}, \dots, \psi_{2r-1,j}$ satisfy the hypotheses of Lemma 3.2.3 whenever $r \le r_j$. Thus we may apply Lemma 3.3.1 to relate $S_{s,r_j}(P, Q_j, R; \boldsymbol{\psi}_j)$ to $T_{s,r_j}(P, Q_j, R, \phi_{j+1}; \boldsymbol{\psi}_j)$. The following lemma then relates $T_{s,r_j}(P, Q_j, R, \phi_{j+1}; \boldsymbol{\psi}_j)$ to $S_{s,r_{j+1}}(P, Q_j, R; \boldsymbol{\psi}_{j+1})$ and hence allows us to repeat the differencing process.

Lemma 3.4.3. *Suppose that $r \le 2w$ and $0 \le j < k$. Then given $\varepsilon > 0$, there exists $\eta = \eta(\varepsilon, s, k)$ such that whenever $R \le P^\eta$ one has*

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

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

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

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

and

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

Then on considering the underlying Diophantine equations, we have

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

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

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

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

and by using Hölder's inequality twice as in [69], we find that

$$\tilde{T}_{s,r} \ll P^{(3-2k\phi_{j+1})r+\varepsilon} \tilde{H}_j^2 \tilde{M}_{j+1}^2 S_s(Q_{j+1}, R). \tag{3.37}$$

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

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

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

$$U_1 \leq \sum_{\boldsymbol{\eta}\in\{\pm 1\}^r} U_2(\boldsymbol{\eta}),$$

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

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

with $\mathbf{z}, \mathbf{w}, \mathbf{u}, \mathbf{v}, \tilde{\mathbf{u}}, \tilde{\mathbf{v}}, \mathbf{h}, \mathbf{g}, \mathbf{m}, \mathbf{n}$ satisfying (3.13), (3.18), (3.19), (3.35), and (3.36), and with

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

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

On writing

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

we have by Hölder's inequality that

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

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

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

and the lemma follows on combining this with (3.37). □

In analogy with Lemma 4.2 of [69], one might hope to refine the above argument to allow the factor of $P^{(3-2k\phi_{j+1})r}$ in the first term of the estimate to be replaced by P^{2r}, but it is not clear that this can be achieved. As will be seen in Section 3.6, such an improvement would have a significant impact on the strength of our repeated efficient differencing procedure.

3.5 Mean Value Estimates Based on Single Differencing

In this section, we consider estimates for $S_s(P, R)$ arising from a single efficient difference, reserving the full power of the preceding analysis for Section 3.6.

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

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

Further, let

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

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

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

and

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

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

Lemma 3.5.1. *Let $\theta = 1/2k$, and suppose that $s \ge k^2/(1-\theta)$. If $\Delta_s \le k(k+1)$ is an admissible exponent, then the exponent $\Delta_{s+r} = \Delta_s(1-\theta)$ is admissible.*

Proof. By Lemmata 3.3.1 and 3.3.3, we have

$$\begin{aligned} S_{s,r}(P,P,R;\boldsymbol{\psi}_0) \ll P^{3r-1}S_s(P,R) &+ P^{(3+\theta)s+2r+\varepsilon} \\ &+ P^\varepsilon M^{4s-2+2r(2r-1)}\tilde{T}_{s,r}(P,P,R,\theta;\boldsymbol{\psi}_0) \end{aligned} \tag{3.38}$$

for γ sufficiently small, and by the argument of the proof of Lemma 3.4.3 we have

$$\tilde{T}_{s,r}(P,P,R,\theta;\boldsymbol{\psi}_0) \ll P^{(3-2k\theta)r+\varepsilon}M^2S_s(Q,R) + \mathcal{M}_{s,r}(P,Q,R). \tag{3.39}$$

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

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

Hence on recalling Lemma 3.2.1 and considering the underlying Diophantine equations, we obtain from (3.38) and (3.39) that

$$\begin{aligned} S_{s+r}(P,R) \ll P^{2s+2r+\varepsilon} + S_{s,r}(P,P,R;\boldsymbol{\psi}_0) &\ll P^{3r-1}S_s(P,R) \\ &+ P^{(3+\theta)s+2r+\varepsilon} + P^{2r+\varepsilon}M^{4s+2r(2r-1)}S_s(Q,R). \end{aligned} \tag{3.40}$$

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

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

where

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

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

and

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

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

Proof of Theorem 3.1. Let s_1 be as in the statement of the theorem, and suppose that $s \geq s_1$. Choose an integer t with $s \equiv t \pmod{r}$ and $s_1 - r < t \leq s_1$. Then since $\Delta_t = k(k+1)$ is trivially admissible, we find by repeated use of Lemma 3.5.1 that the exponent

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

is admissible, and this completes the proof.

3.6 Estimates Arising from Repeated Differencing

In this section, we explore the possibility of obtaining improved mean value estimates by employing our efficient differencing procedure repeatedly. As we take more differences, we must reduce the number of variables taken in a complete interval, so that the difference polynomials $\boldsymbol{\psi}_j$ will satisfy the hypotheses of Lemma 3.2.3. This complicates the recursion for generating admissible exponents and therefore requires some additional notation. Recall the definition of r_j from (3.31), and write

$$\Omega_j = \sum_{2r_j < l \leq k+1} (k - l + 1) = \tfrac{1}{2}(k - 2r_j + 1)(k - 2r_j). \tag{3.41}$$

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

Lemma 3.6.1. *Suppose that $u \geq k(k+1)$ and that $\Delta_u \leq k(k+1)$ is an admissible exponent. For any integer j with $1 \leq j \leq \sqrt{k}$, define $\phi(j, s, J)$ recursively for $s = u + lr$ and $J = j, \dots, 2$ by $\phi(j, s, j) = 1/2k$,*

$$\phi^*(j, s, J-1) = \frac{1}{4k} + \left(\frac{1}{2} + \frac{2\Omega_{J-1} - \Delta_{s-r}}{8kr_{J-1}}\right)\phi(j, s, J), \tag{3.42}$$

and

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

where

$$\Delta_s = \Delta_{s-r}(1 - \theta_s) + r(2k\theta_s - 1) \tag{3.43}$$

and

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

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

Proof. We start by noting that $0 < \theta_s \le 1/2k$ and that θ_s is an increasing function of s. Now let j denote the least integer with $\phi(j, s+r, 1) = \theta_{s+r}$ and write $\phi_J = \phi(j, s+r, J)$. As in the proof of [69], Theorem 6.1, we have $\phi_J < 1/2k$ whenever $J < j$. In particular, it follows that whenever $J < j$ we have $2\Omega_J - \Delta_s < 0$ and $\phi_J = \phi^*(j, s+r, J)$. We claim that $\phi_J \le \phi_{J+1}$ for $J < j$. By (3.42) and the above remarks, this is equivalent to

$$\phi_{J+1}\left(\frac{1}{2} + \frac{\Delta_s - 2\Omega_J}{8kr_J}\right) \ge \frac{1}{4k}, \tag{3.44}$$

and this is immediate when $J = j-1$, since $\Delta_s - 2\Omega_{j-1} > 0$ and $\phi_j = 1/2k$. Assuming the claim holds for J, then we see from (3.42) that

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

and it follows on using (3.41) that

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

Since $\Delta_s \le k(k+1)$ and $r_J \le r_{J-1}$, we see that (3.44) holds with J replaced by $J-1$, and our claim follows.

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

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

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

$$\tilde{T}_{s,r_J}(P, Q_J, R, \phi_{J+1}; \boldsymbol{\psi}_J) \ll P^{(3-2k\phi_{J+1})r_J+\varepsilon}\tilde{H}_J^2\tilde{M}_{J+1}^2 Q_{J+1}^{\lambda_s} \tag{3.45}$$

for $J = j-1, \ldots, 0$. By Lemma 3.4.3 with j replaced by $j-1$, $r = r_{j-1}$ and $w = r_j$, we have that

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

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

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

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

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

on using the outer induction hypothesis. Thus (3.45) holds in the case $J = j-1$.

Now suppose that (3.45) holds for J. Then, for γ sufficiently small, we have by Lemmata 3.3.1 and 3.3.3 that

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

where

$$\Lambda_1 = 3r_J - 1 + \lambda_s(1 - \phi_1 - \cdots - \phi_J), \tag{3.46}$$

$$\Lambda_2 = 3s(1 - \phi_1 - \cdots - \phi_J) + s\phi_{J+1} + 2r_J, \tag{3.47}$$

and

$$\Lambda_3 = (4s + 2r_J(2r_J - 1))\phi_{J+1} + (3 - 2k\phi_{J+1})r_J + \lambda_s(1 - \phi_1 - \cdots - \phi_{J+1}). \tag{3.48}$$

Now since $J \le \sqrt{k}$, we have $r_J \sim k/2$, and it follows easily that $\Lambda_1 \le \Lambda_3$ and $\Lambda_2 \le \Lambda_3$ for $s \ge k(k+1)$ and k sufficiently large. Hence by Lemma 3.4.3 we have

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

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

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

where

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

By (3.41) and (3.42), we have

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

and hence

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

since $\phi_{J+1} \geq \phi_J$ and $\phi_J \leq 1/2k$. Thus (3.45) holds with J replaced by $J-1$, so this completes the inner induction. Now we apply (3.45) with $J = 0$ to obtain

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

whence by Lemmata 3.2.1, 3.3.1, and 3.3.3 we have (for γ sufficiently small) that

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

where Λ_1, Λ_2, and Λ_3 are given by (3.46), (3.47), and (3.48) with $J = 0$. Therefore, the exponent

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

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

Next we investigate the size of the admissible exponents supplied by Lemma 3.6.1.

Lemma 3.6.2. *Suppose that* $s > k(k+1)+r$ *and that* Δ_{s-r} *is an admissible exponent satisfying*

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

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

$$\delta_s + \log \delta_s = \delta_{s-r} + \log \delta_{s-r} - \frac{3}{4k} + \frac{1}{k(\log k)^{3/2}}. \tag{3.49}$$

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

Proof. The proof is nearly identical to that of [69], Lemma 6.2. In view of (3.43), we may assume that $0 \leq \Delta_s \leq 2rk$ and hence that $0 \leq \delta_s \leq \frac{1}{2}$. By Lemma 3.6.1 with

$$j = \left[\frac{1}{2}(\log k)^{1/4}\right] + 1, \tag{3.50}$$

we see that the exponent

$$\Delta_s = \Delta_{s-r}(1 - \theta) + r(2k\theta - 1) = 4kr\delta_{s-r} - r + 2rk(1 - 2\delta_{s-r})\theta, \tag{3.51}$$

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

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

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

$$\phi_{J-1} \le \frac{1}{4k} + \frac{1}{2}(1-\delta')\phi_J, \tag{3.52}$$

where

$$\delta' = \frac{\Delta_{s-r} - (\log k)^{1/2}}{4kr} \ge \delta_{s-r}(1-(\log k)^{-3/2}). \tag{3.53}$$

An easy induction using (3.52) shows that

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

and therefore

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

Write $L = (\log k)^{-3/2}$. Since the expression on the right hand side of the above inequality is a decreasing function of δ', we see from (3.50) and (3.53) that

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

for k sufficiently large. It now follows from (3.51) that

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

where

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

Hence if δ_s is defined by (3.49), then since $\log(1-x) \le -x$ for $0 < x < 1$, we have

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

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

We are now fully equipped to prove Theorem 3.2.

Proof of Theorem 3.2. We first note that the theorem is trivial when $1 \leq s \leq s_0$. Now when $s > s_0$, define δ_s to be the unique positive solution of the equation

$$\delta_s + \log \delta_s = 1 - \frac{3(s-s_0)}{4rk} + \frac{s-s_0}{rk(\log k)^{3/2}}. \tag{3.54}$$

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

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

is trivially admissible. Then we have

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

and hence

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

for $k \geq 2$. It it follows that the exponent $4rk\delta_s$ is admissible, since $\delta + \log \delta$ is an increasing function of δ. Now suppose that $\Delta_{s-r} = 4kr\delta_{s-r}$ is admissible, where $s_0 + r < s \leq s_1$. We have by (3.54) that $\delta_{s-r} \leq 1$ and

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

from which it follows that

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

Thus Lemma 3.6.2 shows that $\Delta_s = 4rk\gamma_s$ is admissible, where γ_s is the unique positive solution of

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

Applying (3.54) with s replaced by $s - r$ now shows that $\gamma_s + \log \gamma_s = \delta_s + \log \delta_s$, whence $\gamma_s = \delta_s$, and the induction is complete.

The theorem now follows immediately in the case where $1 \leq s \leq s_1$, since from (3.54) and the definition of s_1 we see that

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

for k sufficiently large.

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

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

is admissible, and the theorem follows on applying Lemma 3.5.1 repeatedly.

We note that in the presence of the refined version of Lemma 3.4.3 discussed at the end of Section 3.4, we could replace the factor of r in the second term of (3.43) by $2r$ and the $3/4k$ term in (3.49) by $1/k$. Hence we would obtain admissible exponents that decay like k^2e^{-2s/k^2} in many cases of interest.

CHAPTER IV

Applications of the Mean Value Theorems

4.1 Overview

The technical apparatus developed in the previous chapter, culminating in the mean value estimates of Theorems 3.1 and 3.2, allows us to consider a variety of Diophantine problems that would not otherwise be accessible. In this chapter, we are primarily concerned with problems involving forms of large degree, while the application of our methods to cubic forms is illustrated in Chapter 5.

In applications involving the circle method, the mean value estimates of Chapter 3 are of fundamental importance to the treatment of the minor arcs, but we also require Weyl estimates in order to proceed with the analysis. Fortunately, our mean value estimates themselves give rise to Weyl estimates by way of the large sieve inequality. Thus in Section 4.2 we will be able to prove

Theorem 4.1. *For $\mu > 0$, define $\mathfrak{m}_\mu$ to be the set of $\boldsymbol{\alpha} \in \mathbb{R}^{k+1}$ such that whenever $a_i \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy*

$$(a_0, \dots, a_k, q) = 1 \quad \text{and} \quad |q\alpha_i - a_i| \le P^{\mu-k}R^k \quad (0 \le i \le k)$$

one has $q > P^\mu R^{k+1}$. Suppose that $0 < \lambda \le \frac{1}{2}$ and that Δ_s denotes an admissible exponent. Then given $\varepsilon > 0$ there exists $\eta = \eta(\varepsilon, k)$ such that whenever $R \le P^\eta$ one has

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

where

$$\sigma(\lambda) = \max_{2s \ge k+1} \frac{\lambda - (1-\lambda)\Delta_s}{2s}. \tag{4.1}$$

The following special case of Theorem 4.1 will be sufficient for most of our applications in this chapter.

Corollary 4.1.1. *Given $\varepsilon > 0$, there exists $\eta = \eta(\varepsilon, k)$ such that whenever $R \le P^\eta$ one has*

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

where

$$\sigma_1(k)^{-1} \sim \tfrac{28}{3} k^3 \log k \quad as \quad k \to \infty.$$

The estimates contained in Theorem 4.1 have immediate applications to the problem of giving localized bounds for the fractional parts of polynomials in two variables. Thus in Section 4.3 we are able to establish the following result.

Theorem 4.2. *Suppose that $\boldsymbol{\alpha} \in \mathbb{R}^{k+1}$, and let Δ_s denote an admissible exponent. Then given $\varepsilon > 0$, there exists $N_0 = N_0(\varepsilon, k)$ such that whenever $N > N_0$ one has*

$$\min_{1 \le m,n \le N} ||\alpha_0 m^k + \alpha_1 m^{k-1} n + \cdots + \alpha_k n^k|| < N^{\varepsilon - \tau(k)},$$

where

$$\tau(k) = \max_{2s \ge k+1} \frac{1 - k\Delta_s}{2s(k+1) + 1 + \Delta_s}.$$

In particular, we have

Corollary 4.2.1. *Given $\boldsymbol{\alpha} \in \mathbb{R}^{k+1}$ and $\varepsilon > 0$, there exists $N_0 = N_0(\varepsilon, k)$ such that whenever $N > N_0$ one has*

$$\min_{1 \le m,n \le N} ||\alpha_0 m^k + \alpha_1 m^{k-1} n + \cdots + \alpha_k n^k|| < N^{\varepsilon - \rho(k)},$$

where

$$\rho(k)^{-1} \sim \tfrac{14}{3} k^3 \log k \quad as \quad k \to \infty.$$

As remarked in the introduction, Theorem 4.2 and Corollary 4.2.1 are primarily of interest when both α_0 and α_k are non-zero, since one may otherwise obtain superior results by specializing one variable and applying single-variable methods.

Our remaining applications require the use of the Hardy-Littlewood method, so in Section 4.4 we take a brief detour to develop the necessary major arc approximations for our exponential sums.

In Section 4.5, we consider the multidimensional analogue of Waring's problem discussed in the introduction. Let $W_s(\mathbf{n}, P)$ denote the number of solutions of the system of equations

$$x_1^{k-j} y_1^j + \cdots + x_s^{k-j} y_s^j = n_j \quad (0 \le j \le k) \tag{4.2}$$

with $x_i, y_i \in [1, P] \cap \mathbb{Z}$. The following result quantifies Theorem 6 by providing a lower bound for $W_s(\mathbf{n}, P)$.

Theorem 4.3. *Suppose that*

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

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

$$\eta_1^{k-j}\xi_1^j + \cdots + \eta_s^{k-j}\xi_s^j = \mu_j \quad (0 \leq j \leq k) \tag{4.3}$$

has a non-singular real solution with $0 < \eta_i, \xi_i < 1$. Suppose also that the system (4.2) *has a non-singular p-adic solution for all primes p. Then there exist positive numbers $\delta = \delta(s, k, \boldsymbol{\mu})$ and $P_0 = P_0(s, k, \boldsymbol{\mu})$ such that, whenever*

$$|n_j - P^k \mu_j| < \delta P^k \quad (0 \leq j \leq k) \tag{4.4}$$

and $P > P_0$, one has

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

Finally, in Section 4.6, we consider the problem of counting rational lines on the hypersurface defined by an additive equation. Let $c_1, \ldots, c_s$ be nonzero integers, and write $N_s(P)$ for the number of solutions of the polynomial equation

$$c_1(x_1 t + y_1)^k + \cdots + c_s(x_s t + y_s)^k = 0 \tag{4.5}$$

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

$$c_1 x_1^{k-j} y_1^j + \cdots + c_s x_s^{k-j} y_s^j = 0 \quad (0 \leq j \leq k) \tag{4.6}$$

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

Theorem 4.4. *Suppose that*

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

and that the system of equations (4.6) *has a non-singular real solution and a non-singular p-adic solution for all primes p. Then for P sufficiently large one has*

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

Given a line $\ell : \mathbf{x}t + \mathbf{y}$, we define the height of ℓ by $h(\ell) = \max(|x_i|, |y_i|)$. Among the solutions counted by $N_s(P)$, we may of course have several that correspond to the same line, so Theorem 4.4 does not directly yield a lower bound for the number of lines on the hypersurface

$$c_1 z_1^k + \cdots + c_s z_s^k = 0 \tag{4.7}$$

with $h(\ell) \leq P$. In Section 4.6, however, we will actually derive the estimate of Theorem 4.4 with the variables restricted to dyadic-type intervals and then show that in this situation the number of solutions of (4.6) corresponding to any particular line is at most $O(1)$. Thus we will prove

Theorem 4.5. *Let $L_s(P)$ denote the number of distinct rational lines ℓ lying on the hypersurface* (4.7) *and satisfying $h(\ell) \leq P$. Then, under the hypotheses of Theorem* 4.4*, one has*

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

It is worth remarking that the p-adic solubility conditions imposed in the above theorems in fact need only be checked for finitely many primes p, as we will see in Sections 4.5 and 4.6 that primes sufficiently large in terms of k are dealt with unconditionally using exponential sums.

Throughout this chapter, any statement involving ε and R is taken to mean that for every $\varepsilon > 0$ there exists a positive number $\eta = \eta(\varepsilon, s, k)$ such that the result holds whenever $R \leq P^\eta$.

4.2 Weyl Estimates

Here we obtain the estimates for smooth Weyl sums quoted in Theorem 4.1 by making simple modifications in the corresponding argument of Wooley [69]. In the end, a standard application of the large sieve inequality shows that these estimates follow from the mean value estimates of Theorems 3.1 and 3.2. Let

$$\mathcal{C}_q(Q) = \{x \in \mathbb{Z} \cap [1, Q] : s_0(x) | s_0(q)\}, \tag{4.8}$$

write

$$\psi(x, y; \boldsymbol{\alpha}) = \sum_{i=0}^{k} \alpha_i x^{k-i} y^i, \tag{4.9}$$

and define the exponential sum

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

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

$$\mathcal{B}(M, \pi, R) = \{v \in \mathbb{N} : M < v \leq M\pi,\ \pi | v,\ \text{and } p | v \Rightarrow \pi \leq p \leq R\}. \tag{4.10}$$

We have the following analogue of [69], Lemma 7.2.

Lemma 4.2.1. *Suppose that* $\boldsymbol{\alpha} \in \mathbb{R}^{k+1}$ *and* $r \in \mathbb{N}$*. Then, whenever*

$$R \le M < Q \ll P \quad \text{and} \quad R \le M' < Q' \ll P,$$

we have

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

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

Proof. By Lemma 10.1 of Vaughan [53], we have

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

where

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

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

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

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

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

and the lemma follows on noting that

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

for $X \geq 1$. $\square$

Theorem 4.1 is an easy consequence of the following lemma.

Lemma 4.2.2. *Suppose that $0 < \lambda \leq \frac{1}{2}$ and write $M = P^\lambda$. Let j be an integer with $0 \leq j \leq k$, and let $\boldsymbol{\alpha} \in \mathbb{R}^{k+1}$. Suppose that $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $(a, q) = 1$, $|q\alpha_j - a| \leq \frac{1}{2}(MR)^{-k}$, $q \leq 2(MR)^k$, and either $|q\alpha_j - a| > MP^{-k}$ or $q > MR$. Then whenever s is a natural number with $2s > \max(j, k-j)$ and the exponent Δ_s is admissible we have*

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

Proof. By Lemma 3.2.4, along with a standard estimate for the divisor function, we see that $\operatorname{card}(\mathcal{C}_q(X)) \ll X^\varepsilon$ whenever $\log q \ll \log X$, and it follows that

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

Thus by Lemma 4.2.1 there exist $d, e \in \mathcal{C}_q(M/R)$, $\theta, \theta' \in [0,1]$ and primes $\pi, \pi' \leq R$ such that

$$f(\boldsymbol{\alpha}; P, R) \ll P^{2+\varepsilon} M^{-1} + P^\varepsilon g(\boldsymbol{\alpha}; d, e, \pi, \pi', \theta, \theta'), \tag{4.11}$$

where

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

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

$(h',q)|d^{k-j}e^j$. In this instance, a simple application of the Chinese Remainder Theorem shows that

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

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

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

$$\sum_{i=1}^{s} u_i^{k-j}(u_i')^j = y_j \quad (0 \le j \le k) \tag{4.12}$$

with

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

and

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

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

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

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

$$|g(\boldsymbol{\alpha})|^{2s} \ll P^{\varepsilon} M^{4s-2} Q^{k^2} \sum_{\mathbf{y}}{}^{*} \sum_{\substack{v\in\mathcal{B}(M/d,\pi,R)\\(v,q)=1}} \sum_{v'\in\mathcal{V}_r} \left| \sum_{y_j} c_{\mathbf{y}} e(\alpha_j (vd)^{k-j}(v'e)^j y_j) \right|^2, \tag{4.13}$$

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

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

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

In particular, if $q > MR/e$, then the elements of $\mathcal{V}_r$ are distinct modulo q, so the $\alpha_j(vd)^{k-j}(v'e)^j$ with $v' \in \mathcal{V}_r$ are spaced at least $\frac{1}{2}q^{-1}$ apart. Thus it suffices to consider the case when v_1' and v_2' are distinct elements of $\mathcal{V}_r$ with $v_1' \equiv v_2' \pmod{q}$ and $q \leq MR/e$. In this case we have

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

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

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

and thus on applying the large sieve inequality (see for example [19]) to (4.13) we obtain

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

But $\sum_{\mathbf{y}} |c_{\mathbf{y}}|^2 \leq S_s(P/M, R)$ and $q \leq 2(MR)^k \ll (P/M)^k$ so on recalling (4.11) we have

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

as required. $\square$

Proof of Theorem 4.1. Suppose that $\boldsymbol{\alpha} \in \mathfrak{m}_{\lambda(k+1)}$ and write $M = P^{\lambda}$. By Dirichlet's Theorem there exist $b_i \in \mathbb{Z}$ and $q_i \in \mathbb{N}$ with $(b_i, q_i) = 1$ such that

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

If for some j we have either

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

then the desired conclusion follows from Lemma 4.2.2. Otherwise, write $q = [q_0, \ldots, q_k]$ and $a_i = b_iq/q_i$. Then $(a_0, \ldots, a_k, q) = 1$, and for each i we have

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

and

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

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

Proof of Corollary 4.1.1. We apply Theorem 4.1 with $\lambda = \frac{1}{2(k+1)}$. By (4.1), we have

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

Then on taking

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

we have by Theorem 3.2 that the exponent

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

is admissible. It follows that

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

We remark that the proof of Lemma 4.2.2, with trivial modifications, may be applied to more general exponential sums of the form

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

provided that $P \asymp Q$, and hence Theorem 4.1 and Corollary 4.1.1 hold in this case as well. This observation will be useful in the analysis of Section 4.6.

4.3 Fractional Parts

The following lemma allows us to deduce information about the distribution of fractional parts from the exponential sum estimates of Theorem 4.1.

Lemma 4.3.1. *Suppose that $g(m,n)$ is a non-negative function, L is a positive integer, and $\alpha_{m,n}$ are real numbers. If*

$$\sum_{l=1}^{L} \left| \sum_{m=1}^{N} \sum_{n=1}^{N} g(m,n) e(l\alpha_{m,n}) \right| < \frac{1}{6} \sum_{m=1}^{N} \sum_{n=1}^{N} g(m,n),$$

then for any β there is a solution of $||\alpha_{m,n} + \beta|| < L^{-1}$ with $1 \leq m, n \leq N$.

Proof. For each integer n with $1 \leq n \leq N^2$, we can write $n = qN + r$ for unique integers $q = q(n)$ and $r = r(n)$ satisfying $0 \leq q \leq N-1$ and $1 \leq r \leq N$. Now we apply Lemma 5 of Harman [26] to the sequence $\alpha'_n = \alpha_{q+1,r}$ $(1 \leq n \leq N^2)$. □

The localized fractional parts estimates of Theorem 4.2 now follow in a routine manner.

Proof of Theorem 4.2. Fix $\boldsymbol{\alpha} \in \mathbb{R}^{k+1}$, let P be a large real number, and let ε be a real number with $0 < \varepsilon < \tau$. Further, write $H_1 = P^{\tau-\varepsilon}$ and

$$T_1(\boldsymbol{\alpha}) = \sum_{1 \le h \le H_1} |f(h\boldsymbol{\alpha}; P, R)|, \tag{4.14}$$

where $R = P^{\eta}$ and $\eta < \eta_0(\varepsilon, k)$. We divide into cases.

(i) Suppose that there exist $h, \mathbf{b}$, and q with $1 \le h \le H_1$, $\mathbf{b} \in \mathbb{Z}^{k+1}$, $q \in \mathbb{N}$, $(q, b_0, \dots, b_k) = 1$, and $q \le P^{1-\tau} R^{k+1}$ such that

$$|qh\alpha_i - b_i| \le P^{1-\tau-k} R^k \quad (0 \le i \le k).$$

Then for each i, we have

$$||\alpha_i (qh)^k|| \le |(qh)^{k-1}(qh\alpha_i - b_i)| < H_1^{k-1} P^{-k\tau} R^{k^2+k-1} < \frac{1}{(k+1)H_1}$$

for η sufficiently small. Thus on noting that $qh \le P$ we obtain

$$\min_{1 \le m,n \le P} ||\psi(m, n; \boldsymbol{\alpha})|| \le ||\psi(qh, qh; \boldsymbol{\alpha})|| \le \sum_{i=0}^{k} ||\alpha_i (qh)^k|| < H_1^{-1},$$

which completes the proof in this case.

(ii) If the hypothesis of case (i) does not hold, then we may apply Theorem 4.1 with $\lambda = \frac{1-\tau}{k+1}$ to obtain

$$\max_{1 \le h \le H_1} |f(h\boldsymbol{\alpha}; P, R)| \ll P^{2-\sigma(\lambda)+\varepsilon/2},$$

and hence

$$T_1(\boldsymbol{\alpha}) \ll P^{2+\tau-\sigma(\lambda)-\varepsilon/2}.$$

Now choose s so that

$$\tau = \frac{1 - k\Delta_s}{2s(k+1) + 1 + \Delta_s}.$$

Then a simple calculation shows that

$$\tau = \frac{1 - \tau - (k+\tau)\Delta_s}{2s(k+1)} \le \sigma(\lambda),$$

whence $T_1(\boldsymbol{\alpha}) = o(P^2)$. The theorem now follows by applying Lemma 4.3.1 with

$$g(m,n) = \begin{cases} 1 & \text{if } m,n \in \mathcal{A}(P,R), \\ 0 & \text{otherwise,} \end{cases}$$

on recalling (see for example [55]) that $\operatorname{card}(\mathcal{A}(P,R)) \gg P$ when $R = P^\eta$.

Proof of Corollary 4.2.1. As in the proof of Corollary 4.1.1, we take

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

and apply Theorem 3.2 to estimate Δ_s. In the notation of Theorem 4.2, we find that

$$\tau(k) \geq \frac{1 + O((\log k)^{-1/3})}{2s(k + O(1))} \sim \left(\frac{14}{3}k^3 \log k\right)^{-1},$$

and this completes the proof.

4.4 Generating Function Asymptotics

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

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

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

and we also define

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

$$v(\boldsymbol{\beta}) = \int_0^P \int_0^P e(\beta_0 \gamma^k + \beta_1 \gamma^{k-1} \nu + \cdots + \beta_k \nu^k)\, d\gamma\, d\nu, \tag{4.15}$$

and

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

Lemma 4.4.1. *When $\alpha_i = a_i/q + \beta_i$ for $0 \le i \le k$, one has*

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

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

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

where

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

Thus on making the change of variables $\gamma = qz + r$ and $\nu = qw + s$ in (4.15), we obtain

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

where

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

Using the mean value theorem, we find that

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

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

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

from which the lemma follows. □

We now begin to analyze the sum $f(\boldsymbol{\alpha}; P, R)$. First we record a partial summation lemma analogous to Lemma 2.6 of Vaughan [55].

Lemma 4.4.2. *Let $c_{m,n}$ be arbitrary complex numbers, and suppose that $F(x,y)$ has continuous partial derivatives on $[0,X] \times [0,Y]$. Then*

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

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

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

and

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

Thus we can write

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

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

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

Lemma 4.4.3. *Let τ be a fixed number, and suppose that $R \leq m, n \leq R^\tau$. Then*

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

Proof. By Lemma 5.3 of Vaughan [53], we have

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

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

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

$$\mathfrak{N}(q,\mathbf{a}) = \{\boldsymbol{\alpha} \in \mathbb{T}^{k+1} : |\alpha_i - a_i/q| \leq WP^{-k}\ (0 \leq i \leq k)\} \tag{4.16}$$

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

$$w(\boldsymbol{\beta}) = \int_R^P \int_R^P \rho\left(\frac{\log\gamma}{\log R}\right)\rho\left(\frac{\log\nu}{\log R}\right)e(\beta_0\gamma^k + \cdots + \beta_k\nu^k)\,d\gamma\,d\nu. \tag{4.17}$$

Lemma 4.4.4. *Suppose that $\boldsymbol{\alpha} \in \mathfrak{N}(q,\mathbf{a})$ with $q \leq R$, and write $\beta_i = \alpha_i - a_i/q$. Then we have*

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

Proof. By arguing as in the proof of Vaughan [53], Lemma 5.4, we obtain

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

whenever $R\le m,n\le P$, and hence by Lemma 4.4.3 we have

$$\begin{aligned}\sum_{\substack{x\in\mathcal{A}(m,R)\\ y\in\mathcal{A}(n,R)}} e\left(\frac{a_0x^k+\cdots+a_ky^k}{q}\right) &= q^{-2}S(q,\mathbf{a})\sum_{\substack{x\in\mathcal{A}(m,R)\\ y\in\mathcal{A}(n,R)}} 1 + O\left(\frac{q^2P^2}{\log P}\right)\\ &= q^{-2}S(q,\mathbf{a})\rho\left(\frac{\log m}{\log R}\right)\rho\left(\frac{\log n}{\log R}\right)mn + E_1,\end{aligned}$$

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

$$c_{x,y} = e\left(\frac{a_0x^k+\cdots+a_ky^k}{q}\right)1_{\mathcal{B}}(x,y) \quad\text{and}\quad F(x,y)=e(\beta_0x^k+\cdots+\beta_ky^k)$$

in Lemma 4.4.2 we find that

$$f(\boldsymbol{\alpha};P,R)=\sum_{1\le x,y\le P} c_{x,y}F(x,y) = S_0 - S_1 + S_2, \tag{4.18}$$

where

$$S_0=\sum_{x,y\in\mathcal{A}(P,R)} e\left(\frac{a_0x^k+\cdots+a_ky^k}{q}\right)\big(e(\beta_0P^k+\cdots+\beta_ky^k)+e(\beta_0x^k+\cdots+\beta_kP^k)\big),$$

$$S_1=\sum_{x,y\in\mathcal{A}(P,R)} e\left(\frac{a_0x^k+\cdots+a_ky^k}{q}\right)e(\beta_0P^k+\cdots+\beta_kP^k),$$

and

$$S_2=\int_0^P\int_0^P \frac{\partial^2}{\partial\gamma\partial\nu}\left(e(\beta_0\gamma^k+\cdots+\beta_k\nu^k)\right)\sum_{\substack{x\in\mathcal{A}(\gamma,R)\\ y\in\mathcal{A}(\nu,R)}} e\left(\frac{a_0x^k+\cdots+a_ky^k}{q}\right)d\nu\,d\gamma.$$

From our observations above, we see immediately that

$$S_1 = q^{-2}S(q,\mathbf{a})P^2\rho(1/\eta)^2e(\beta_0P^k+\cdots+\beta_kP^k)+O\left(\frac{q^2P^2}{\log P}\right). \tag{4.19}$$

We next observe that, by equation (8.13) of Wooley [61], one has

$$\sum_{x\in\mathcal{A}(m,R)} e\left(\frac{a_0x^k+\cdots+a_ky^k}{q}\right) = q^{-1}S(q,\mathbf{a};y)\,m\,\rho\left(\frac{\log m}{\log R}\right)+O\left(\frac{qP}{\log P}\right), \qquad (4.20)$$

where

$$S(q,\mathbf{a};y)=\sum_{1\le x\le q} e\left(\frac{a_0x^k+a_1x^{k-1}y+\cdots+a_ky^k}{q}\right).$$

If we write $S_0=S_3+S_4$, then by (4.20) we have

$$\begin{aligned} S_3 &= \sum_{x,y\in\mathcal{A}(P,R)} e\left(\frac{a_0x^k+\cdots+a_ky^k}{q}\right)e(\beta_0P^k+\cdots+\beta_ky^k)\\ &= q^{-1}\rho(1/\eta)P\sum_{y\in\mathcal{A}(P,R)} S(q,\mathbf{a};y)e(\beta_0P^k+\cdots+\beta_ky^k)+O\left(\frac{qP^2}{\log P}\right), \end{aligned}$$

and then by partial summation

$$\begin{aligned} S_3 &= q^{-1}P\,\rho(1/\eta)\,T(P)\,e(\beta_0P^k+\cdots+\beta_kP^k)\\ &\quad -\,q^{-1}P\,\rho(1/\eta)\int_R^P T(\nu)\frac{\partial}{\partial\nu}\left(e(\beta_0P^k+\cdots+\beta_k\nu^k)\right)d\nu+O\left(\frac{qP^2}{\log P}\right), \end{aligned}$$

where

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

But on using the obvious analogue of (4.20) we find that

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

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

$$\frac{\partial}{\partial\nu}\left(e(\beta_0P^k+\cdots+\beta_k\nu^k)\right)\ll W/P.$$

Therefore we obtain

$$S_3=Q\,\rho(1/\eta)\,P\,e(\beta_0P^k+\cdots+\beta_kP^k)-Q\,I(P)+O\left(\frac{qP^2W}{\log P}\right), \qquad (4.21)$$

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

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

Integration by parts yields

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

but

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

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

$$\begin{aligned}I(\gamma) &= \rho(1/\eta)P\, e(\beta_0\gamma^k + \cdots + \beta_k P^k) \\ &\qquad - \int_R^P e(\beta_0\gamma^k + \cdots + \beta_k\nu^k)\, \rho\left(\frac{\log\nu}{\log R}\right) d\nu + E_2(\gamma),\end{aligned} \tag{4.22}$$

where $E_2(\gamma) \ll P/\log P$, so it follows from (4.21) that

$$S_3 = Q\int_R^P \rho\left(\frac{\log\nu}{\log R}\right) e(\beta_0 P^k + \cdots + \beta_k\nu^k)\, d\nu + O\left(\frac{qP^2W}{\log P}\right). \tag{4.23}$$

Moreover, an identical argument shows that

$$S_4 = Q\int_R^P \rho\left(\frac{\log\gamma}{\log R}\right) e(\beta_0\gamma^k + \cdots + \beta_k P^k)\, d\gamma + O\left(\frac{qP^2W}{\log P}\right). \tag{4.24}$$

We now deal with S_2. A simple calculation shows that

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

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

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

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

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

and on integrating by parts we get

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

Then from (4.22) we finally obtain

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

where

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

and the lemma follows on recalling (4.18), (4.19), (4.23), and (4.24). $\square$

4.5 A Multidimensional Analogue of Waring's Problem

Here we establish Theorem 4.3 by a fairly straightforward application of the Hardy-Littlewood method. Let P be a large positive number, and put $R = P^\eta$, where $\eta \le \eta_0(\varepsilon, k)$. Let $F(\boldsymbol{\alpha})$ be as in the previous section, and write $f(\boldsymbol{\alpha}) = f(\boldsymbol{\alpha}; P, R)$. Further, put $s = t + 2u + v$, and let

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

Then we have $W_s(\mathbf{n}, P) \ge R_s(\mathbf{n})$, so it suffices to obtain a lower bound for $R_s(\mathbf{n})$. We dissect $\mathbb{T}^{k+1}$ into major and minor arcs as follows. Recalling the notation of Theorem 4.1, define

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

We take

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

where Δ_u is as in Theorem 3.2 and $\sigma_1(k)$ is as in Corollary 4.1.1. A simple calculation shows that $v \ll k^2$, and hence

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

On applying the aforementioned theorem and corollary, we find that

$$\begin{aligned}\int_{\mathfrak{m}} |F(\boldsymbol{\alpha})|^t |f(\boldsymbol{\alpha})|^{2u+v} d\boldsymbol{\alpha} &\ll P^{2t} \sup_{\boldsymbol{\alpha}\in\mathfrak{m}} |f(\boldsymbol{\alpha})|^v \int_{\mathbb{T}^{k+1}} |f(\boldsymbol{\alpha})|^{2u} d\boldsymbol{\alpha} \\ &\ll P^{2s-k(k+1)-\delta} \end{aligned} \tag{4.25}$$

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

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

$$\mathfrak{M}(q, \mathbf{a}) = \{\boldsymbol{\alpha} \in \mathbb{T}^{k+1} : |q\alpha_i - a_i| \le P^{1/2-k} R^k \ (0 \le i \le k)\}, \tag{4.26}$$

so that

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

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

Lemma 4.5.1. *Suppose that $\boldsymbol{\alpha} \in \mathfrak{M}(q, \mathbf{a})$, and write $\beta_i = \alpha_i - a_i/q$. Then one has*

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

Proof. This follows immediately from Lemma 4.4.1, together with (4.26). □

The following estimates for $S(q, \mathbf{a})$, $v(\boldsymbol{\beta})$, and $w(\boldsymbol{\beta})$ are essentially immediate from the work of Arkhipov, Karatsuba, and Chubarikov [3].

Lemma 4.5.2. *Whenever $(q, a_0, \ldots, a_k) = 1$, we have*

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

Proof. This follows easily from [3], Lemma II.8, on recalling standard divisor function estimates. □

Lemma 4.5.3. *One has*

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

and

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

Proof. The first estimate follows from [3], Lemma II.2, on making a change of variable, and the second follows in a similar manner (see the comment in the proof of [61], Lemma 8.6) on noting that $\rho(\log\gamma/\log R) \asymp 1$ and is decreasing for $R \leq \gamma \leq P$. $\square$

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

$$\mathfrak{N} = \bigcup_{\substack{1\leq a_0,\ldots,a_k\leq q\leq W\\(q,a_0,\ldots,a_k)=1}} \mathfrak{N}(q,\mathbf{a}).$$

We have the following result, which is closely analogous to [61], Lemma 9.2.

Lemma 4.5.4. *If t is an integer with $t \geq (k+1)^2$, then one has*

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

and

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

for some $\sigma > 0$.

Proof. When $\boldsymbol{\alpha} \in \mathfrak{M}(q,\mathbf{a})$, we have by Lemma 4.5.1 that

$$F(\boldsymbol{\alpha})^t - V(\boldsymbol{\alpha})^t \ll \left(P^{(3/2+\varepsilon)}\right)^t + P^{3/2+\varepsilon}|V(\boldsymbol{\alpha})|^{t-1}, \tag{4.27}$$

and the proof now follows the argument of Wooley [61], Lemma 9.2, employing Lemma 4.5.2 together with the estimate

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

which is immediate from Lemma 4.5.3. $\square$

On making a trivial estimate for $f(\boldsymbol{\alpha})$, it follows directly from Lemma 4.5.4 that

$$\int_{\mathfrak{M}\setminus\mathfrak{N}} |F(\boldsymbol{\alpha})|^t |f(\boldsymbol{\alpha})|^{2u+v} \, d\boldsymbol{\alpha} \ll W^{-\sigma}P^{2s-k(k+1)}, \tag{4.28}$$

for some $\sigma > 0$, so it suffices to deal with the pruned major arcs $\mathfrak{N}$. When $\boldsymbol{\alpha} \in \mathfrak{N}(q,\mathbf{a})$, we have by Lemma 4.4.4 that

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

where

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

On combining this with (4.27) and recalling the definition of $\mathfrak{N}$, we find that

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

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

Now let

$$S(q) = \sum_{\substack{1\le a_0,\dots,a_k\le q\\ (q,a_0,\dots,a_k)=1}} (q^{-2}S(q,\mathbf{a}))^s e\left(\frac{-a_0n_0 - \dots - a_kn_k}{q}\right),$$

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

and

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

Notice that by Lemma 4.5.2 we have $S(q) \ll q^{k+1-s/k+\varepsilon}$, whence

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

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

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

where

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

and put

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

Then when $s \ge (k+1)^2$, it follows easily from Lemmata 4.5.2 and 4.5.3 that

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

and

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

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

$$\int_{\mathfrak{N}} F(\boldsymbol{\alpha})^t f(\boldsymbol{\alpha})^{2u+v} e(-\boldsymbol{\alpha}\cdot\mathbf{n})\, d\boldsymbol{\alpha} = \mathfrak{S}(\mathbf{n})J(\mathbf{n}) + O(P^{2s-k(k+1)}(\log P)^{-\delta}) \tag{4.29}$$

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

Lemma 4.5.5. *Suppose that $s\ge(k+1)^2$, and fix real numbers $\mu_0,\dots,\mu_k$ with the property that the system (4.3) has a non-singular real solution with $0<\eta_i,\xi_i<1$. Then there exists a positive number $\delta'=\delta'(s,k,\boldsymbol{\mu})$ such that, whenever*

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

and P is sufficiently large, one has

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

Proof. After a change of variables, we have

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

where

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

$$T(\boldsymbol{\gamma},\boldsymbol{\nu}) = \prod_{i=t+1}^{s}\rho\left(\frac{\log P\gamma_i}{\log R}\right)\rho\left(\frac{\log P\nu_i}{\log R}\right), \tag{4.30}$$

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

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

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

so that

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

where $\mathcal{C} \subset \mathbb{R}^{k+1}$. Since $(\boldsymbol{\eta},\boldsymbol{\xi}) \in \mathcal{B}$, we see that $\mathcal{C}$ contains a neighborhood of $(\delta_0,\ldots,\delta_k)$ and hence contains the origin when δ' is sufficiently small. Thus after $k+1$ applications of Fourier's Integral Theorem (see for example Davenport [18]) we obtain

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

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

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

Lemma 4.5.6. *The function $S(q)$ is multiplicative.*

Proof. By [3], Lemma II.4, one has $S(qr,\mathbf{a}) = S(q, r^{k-1}\mathbf{a})S(r, q^{k-1}\mathbf{a})$ whenever $(q,r) = 1$, and the result now follows by a standard argument. □

For each prime p, write

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

Whenever $s \ge (k+1)^2$ one finds using Lemmata 4.5.2 and 4.5.6 that

$$\mathfrak{S}(\mathbf{n}) = \prod_p \sigma(p) \tag{4.31}$$

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

$$\frac{1}{2} \le \prod_{p>C(k)} \sigma(p) \le \frac{3}{2}. \tag{4.32}$$

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

$$x_1^{k-j}y_1^j + \cdots + x_s^{k-j}y_s^j \equiv n_j \pmod q \quad (0 \le j \le k). \tag{4.33}$$

Lemma 4.5.7. *One has*

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

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

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

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

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

and the result follows. □

We therefore have

$$\sigma(p) = \lim_{h\to\infty} p^{h(k+1-2s)} M_{\mathbf{n}}(p^h), \tag{4.34}$$

so to show that $\mathfrak{S}(\mathbf{n}) \gg 1$ it suffices to obtain a suitable lower bound for $M_{\mathbf{n}}(p^h)$. In order to deduce this from the existence of non-singular p-adic solutions to (4.2), we need a version of Hensel's Lemma. In what follows, we write $|\cdot|_p$ for the usual p-adic valuation, normalized so that $|p|_p = p^{-1}$.

Lemma 4.5.8. *Let $\psi_1, \ldots, \psi_r$ be polynomials in $\mathbb{Z}_p[x_1, \ldots, x_r]$ with Jacobian $\Delta(\boldsymbol{\psi}; \mathbf{x})$, and suppose that $\mathbf{a} \in \mathbb{Z}_p^r$ satisfies*

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

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

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

Proof. This is Proposition 5.20 of Greenberg [23] with $R = \mathbb{Z}_p$. □

Lemma 4.5.9. *Suppose that the system (4.2) has a non-singular p-adic solution. Then there exists an integer $u = u(p)$ such that whenever $h \ge u$ one has*

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

Proof. We relabel the variables by writing

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

and let $\mathbf{a} = (a_1, \dots, a_{2s})$ be a non-singular p-adic solution of (4.2). Then there exist indices $i_0, \dots, i_k$ such that $\Delta(\boldsymbol{\psi}; a_{i_0}, \dots, a_{i_k}) \neq 0$, so we can find an integer u such that

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

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

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

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

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

and hence

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

Now if $h \geq u$ then there are $p^{(h-u)(2s-k-1)}$ possible choices for the w_i modulo p^h. Moreover, for any fixed choice we may regard ψ_j as a polynomial in the $k+1$ variables $z_{i_0}, \dots, z_{i_k}$ after substituting $z_i = w_i$ on the remaining indices. Thus for each admissible choice of $\mathbf{w}$ we may apply Lemma 4.5.8 to obtain integers $b_{i_0}, \dots, b_{i_k}$ such that $\psi_j(\mathbf{b}; \mathbf{w}) \equiv 0 \pmod{p^h}$ for each j, whence the lemma follows. $\square$

Now by (4.34) and Lemma 4.5.9 we have $\sigma(p) \geq p^{u(k+1-2s)}$ for all primes p, so on combining this with (4.31) and (4.32) we see that $\mathfrak{S}(\mathbf{n}) \gg 1$. Hence the proof of Theorem 4.3 is complete upon recalling Lemma 4.5.5, together with (4.25), (4.28), and (4.29).

4.6 Lines on Additive Equations

We now establish Theorems 4.4 and 4.5 by proceeding much as in the previous section. Before embarking on the circle method, however, we need to make some preliminary observations.

Lemma 4.6.1. *Suppose that* $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{2s}$ *is a solution of* (4.6)*, and let* $a, b, c,$ *and* d *be arbitrary real numbers. Then* $(a\mathbf{x} + b\mathbf{y}, c\mathbf{x} + d\mathbf{y})$ *is also a solution.*

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

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

Then by the binomial theorem we have for each j that

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

and the lemma follows. □

Lemma 4.6.2. *Suppose that the system of equations* (4.6) *has a non-singular real solution* $(\boldsymbol{\eta}, \boldsymbol{\xi})$. *Then we can find a non-singular real solution* $(\boldsymbol{\eta}', \boldsymbol{\xi}')$ *such that* η_i' *and* ξ_i' *are nonzero for each* i.

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

$$\psi_j(\mathbf{x}, \mathbf{y}) = c_1 x_1^{k-j} y_1^j + \cdots + c_s x_s^{k-j} y_s^j,$$

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

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

Hence by using the Implicit Function Theorem as in the proof of [61], Lemma 6.2, we see that there exists a $(2s-k-1)$-dimensional neighborhood T_0 of $(\gamma_{k+1}, \ldots, \gamma_{2s-1})$ and a function $\phi : T_0 \to \mathbb{R}^{k+1}$ such that $\boldsymbol{\gamma} = (\phi(\mathbf{w}), \mathbf{w})$ is a solution of (4.6) whenever $\mathbf{w} \in T_0$. Thus by choosing $\mathbf{w}$ with $|w_i - \gamma_i|$ sufficiently small for $k+1 \leq i \leq 2s-1$, we may assume that $\boldsymbol{\gamma}$ is a non-singular solution whose last $2s-k-1$ coordinates are nonzero. Moreover, a simple calculation shows that at most two of the remaining η_i and at most two of the remaining ξ_i are zero and that either η_i or ξ_i is nonzero for every i. In particular, when $s \geq 5$, there is some i for which $\eta_i \xi_i \neq 0$. Now let

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

and take $b' < \frac{1}{2}b$ and $c' < \frac{1}{2}c$. Then by Lemma 4.6.1 we see that $(\boldsymbol{\eta}', \boldsymbol{\xi}')$ is a solution of (4.6), where $\boldsymbol{\eta}' = \boldsymbol{\eta} + b'\boldsymbol{\xi}$ and $\boldsymbol{\xi}' = c'\boldsymbol{\eta} + \boldsymbol{\xi}$, and it is easy to check that η_i' and ξ_i' are nonzero for each i. The non-singularity follows by continuity on choosing b' and c' sufficiently small. □

By Lemma 4.6.2 we may henceforth suppose that the system (4.6) has a nonsingular real solution $(\boldsymbol{\eta}, \boldsymbol{\xi})$ with η_i and ξ_i nonzero for all i, and by homogeneity we can re-scale to ensure that $0 < |\eta_i|, |\xi_i| < \frac{1}{2}$. For each i, write

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

and

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

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

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

and

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

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

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

Finally, let

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

Then we have $N_s(P) \geq R_s(P)$, so to prove Theorem 4.4 it suffices to obtain a lower bound for $R_s(P)$. We dissect $\mathbb{T}^{k+1}$ into major and minor arcs as follows. Write $c = \max |c_i|$ and $X = cP^{1/2} R^{k+1}$, and define

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

where

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

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

Lemma 4.6.3. *Whenever* $\boldsymbol{\alpha} \in \mathfrak{m}$, *one has* $c_i\boldsymbol{\alpha} \in \mathbf{m}_{1/2}$. *Moreover,*

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

where $\sigma_1(k)$ *is as in Corollary* 4.1.1.

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

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

so on writing

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

we see that

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

so we must have $cq \ge q' > cP^{1/2}R^{k+1}$ and hence $q > P^{1/2}R^{k+1}$. Thus $c_i\boldsymbol{\alpha} \in \mathbf{m}_{1/2}$. The second assertion now follows on recalling the remark at the end of Section 4.2 and noting that we may replace α_j by $-\alpha_j$ as needed so that our sums are over positive integers. □

As in the previous section, we take

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

where Δ_u is as in Theorem 3.2 and $\sigma_1(k)$ is as in Corollary 4.1.1. Then by Hölder's inequality and a change of variables we obtain

$$\begin{aligned}\int_{\mathfrak{m}} |\mathcal{F}(\boldsymbol{\alpha})\mathcal{G}(\boldsymbol{\alpha})|\,d\boldsymbol{\alpha} &\ll P^{2t+v(2-\sigma_1+\varepsilon)} \prod_{i=t+1}^{t+2u} \left(\int_{\mathbb{T}^{k+1}} |f_i(\boldsymbol{\alpha})|^{2u} d\boldsymbol{\alpha}\right)^{1/2u} \\ &\ll P^{2s-k(k+1)-\delta} \end{aligned} \tag{4.35}$$

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

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

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

and

$$V_i(\boldsymbol{\alpha}) = q^{-2}S_i(q,\mathbf{a})v_i(\boldsymbol{\alpha}-\mathbf{a}/q) \qquad (\boldsymbol{\alpha}\in\mathfrak{M}(q,\mathbf{a})).$$

Further, we define the pruned major arcs $\mathfrak{N}$ exactly as in the previous section, again with W a suitable power of $\log P$, and write

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

and

$$W_i(\boldsymbol{\alpha}) = q^{-2}S_i(q,\mathbf{a})w_i(\boldsymbol{\alpha}-\mathbf{a}/q) \qquad (\boldsymbol{\alpha}\in\mathfrak{N}(q,\mathbf{a})).$$

The next several lemmas are simple adaptations of the corresponding results in the previous section.

Lemma 4.6.4. *When $\boldsymbol{\alpha}\in\mathfrak{M}(q,\mathbf{a})$, one has*

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

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

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

Proof. These estimates follow by making trivial modifications in the arguments of Lemmata 4.4.1 and 4.4.4, respectively. □

Lemma 4.6.5. *Whenever $(q,a_0,\dots,a_k)=1$, we have*

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

Proof. Put $d_i=(q,c_i)$. Then by Lemma 4.5.2 we have

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

as required. □

Lemma 4.6.6. *One has*

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

and

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

Proof. The argument is identical to the proof of Lemma 4.5.3. □

Lemma 4.6.7. *If t is an integer with $t \geq (k+1)^2$, then one has*

$$\int_{\mathfrak{M}} |F_i(\boldsymbol{\alpha})|^t \, d\boldsymbol{\alpha} \ll P^{2t-k(k+1)} \tag{4.36}$$

and

$$\int_{\mathfrak{M}\setminus\mathfrak{N}} |F_i(\boldsymbol{\alpha})|^t \, d\boldsymbol{\alpha} \ll W^{-\sigma} P^{2t-k(k+1)} \tag{4.37}$$

for some $\sigma > 0$.

Proof. The result follows as in Lemma 4.5.4 on using Lemmata 4.6.4, 4.6.5, and 4.6.6 in place of the corresponding results in the previous section. □

Once again, Lemma 4.6.7, together with (4.35), allows us to focus attention on the pruned major arcs $\mathfrak{N}$. Let

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

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

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

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

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

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

Then when $s \geq (k+1)^2$, we have by Lemmata 4.6.5 and 4.6.6 that $J \ll P^{2s-k(k+1)}$ and

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

for some $\delta > 0$. Thus, by employing standard arguments based on Lemmata 4.6.4, 4.6.5, and 4.6.6, we obtain

$$\int_{\mathfrak{N}} \mathcal{F}(\boldsymbol{\alpha}) \mathcal{G}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} = \mathfrak{S} J + O(P^{2s-k(k+1)} (\log P)^{-\delta}) \tag{4.38}$$

for some $\delta > 0$.

Lemma 4.6.8. *Whenever $s \geq (k+1)^2$ and P is sufficiently large, one has*

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

Proof. By a change of variables, we find that

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

where

$$\mathcal{B} = [\eta_1^-, \eta_1^+] \times \cdots \times [\eta_s^-, \eta_s^+] \times [\xi_1^-, \xi_1^+] \times \cdots \times [\xi_s^-, \xi_s^+],$$

$$\phi_j(\boldsymbol{\gamma}, \boldsymbol{\nu}) = c_1 \gamma_1^{k-j} \nu_1^j + \cdots + c_s \gamma_s^{k-j} \nu_s^j,$$

and where $T(\boldsymbol{\gamma}, \boldsymbol{\nu})$ is as in (4.30). Now let

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

so that

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

where $\mathcal{C} \subset \mathbb{R}^{k+1}$. Since $(\boldsymbol{\eta}, \boldsymbol{\xi}) \in \mathcal{B}$, we see that $\mathcal{C}$ contains a neighborhood of the origin, whence after $k+1$ applications of Fourier's Integral Theorem we obtain

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

and the result follows as in the proof of Lemma 4.5.5. □

Lemma 4.6.9. *The function $S(q)$ is multiplicative.*

Proof. This is identical to the proof of Lemma 4.5.6. □

Whenever $s \geq (k+1)^2$ one finds using Lemma 4.6.9 that

$$\mathfrak{S} = \prod_p \sigma(p), \quad \text{where} \quad \sigma(p) = \sum_{h=0}^{\infty} S(p^h),$$

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

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

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

$$c_1 x_1^{k-j} y_1^j + \cdots + c_s x_s^{k-j} y_s^j \equiv 0 \pmod{q} \quad (0 \leq j \leq k).$$

Lemma 4.6.10. *One has*

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

Proof. This is identical to the proof of Lemma 4.5.7. □

It follows that

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

so again to show that $\mathfrak{S} \gg 1$ it suffices to show that

$$M(p^h) \geq p^{(h-u)(2s-k-1)}$$

for $p \leq C(k)$, and this follows exactly as in the argument of Lemma 4.5.9. Hence the proof of Theorem 4.4 is complete on assembling (4.35), (4.37), and (4.38) and recalling Lemma 4.6.8.

In order to deduce Theorem 4.5, we need some additional observations.

Lemma 4.6.11. *Let* $(\mathbf{x}, \mathbf{y}), (\mathbf{x}', \mathbf{y}') \in \mathbb{Z}^{2s}$ *be such that* $(x_1, \ldots, x_s) = 1$. *Then* $\mathbf{x}t + \mathbf{y}$ *and* $\mathbf{x}'t + \mathbf{y}'$ *parameterize the same line if and only if*

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

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

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

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

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

Now let $R_s(P, d)$ denote the number of solutions of (4.6) counted by $R_s(P)$ for which $(x_1, \ldots, x_s) = d$. The following estimate will be useful when d is large.

Lemma 4.6.12. *One has*

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

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

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

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

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

The lemma now follows by dissecting $\mathbb{T}^{k+1}$ into major and minor arcs and using (4.35) and (4.36). $\square$

We can now complete the proof of Theorem 4.5.

Proof of Theorem 4.5. Define an equivalence relation on the set of solutions to (4.6) by writing $(\mathbf{x}, \mathbf{y}) \sim (\mathbf{x}', \mathbf{y}')$ whenever $\mathbf{x}t + \mathbf{y}$ and $\mathbf{x}'t + \mathbf{y}'$ define the same line. We need a lower bound for the number of equivalence classes, so by Lemma 4.6.11 it suffices to estimate the number of solutions of (4.6) for which

$$(x_1, \ldots, x_s) = 1 \quad \text{and} \quad 1 \le y_1 \le |x_1|.$$

Let $N_s(P, Q, d)$ be the number of solutions of (4.6) with $\mathbf{x} \in \mathcal{B}P$, $\mathbf{y} \in \mathcal{C}Q$, and $(x_1, \ldots, x_s) = d$, where

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

Then the solutions counted by $R_s(P, d)$ are in bijective correspondence with a subset of those counted by $N_s(P/d, P, 1)$. Moreover, Lemma 4.6.11 shows that two solutions $(\mathbf{x}, \mathbf{y})$ and $(\mathbf{x}', \mathbf{y}')$ counted by $N_s(P/d, P, 1)$ are equivalent if and only if $\mathbf{x} = \mathbf{x}'$ and $\mathbf{y} - \mathbf{y}' = r\mathbf{x}$ for some integer r. Then since

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

where $\tau = 2/|\eta_1|$, we see that each equivalence class contains at most τd members counted by $N_s(P/d, P, 1)$. Thus we see that

$$R_s(P, d) \le N_s(P/d, P, 1) \le \tau d \, L_s(P),$$

and hence for any D we have that

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

Thus by Lemma 4.6.12 there exist positive constants γ_1 and γ_2 such that

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

for P sufficiently large, and hence we have

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

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

CHAPTER V

The Density of Rational Lines on Cubic Hypersurfaces

5.1 Overview

In this chapter, we refine the analysis of Chapters 3 and 4 in the case $k = 3$ by providing a more detailed consideration of the lower moments of the relevant exponential sums. In particular, we are able to establish diagonal behavior for the mean value

$$S(P) = \int_{\mathbb{T}^4} |F(\boldsymbol{\alpha})|^6 d\boldsymbol{\alpha},$$

where

$$F(\boldsymbol{\alpha}) = \sum_{1 \le x, y \le P} e(\alpha_0 x^3 + \alpha_1 x^2 y + \alpha_2 x y^2 + \alpha_3 y^3),$$

and this leads, via the iterative method, to improved estimates for higher moments.

While the assertion of diagonal behavior amounts to the estimate $S(P) \ll P^6$, we can actually establish a more precise result. Notice that, by orthogonality, $S(P)$ is the number of solutions of the system of equations

$$\begin{aligned} x_1^3 + x_2^3 + x_3^3 &= x_4^3 + x_5^3 + x_6^3 \\ x_1^2 y_1 + x_2^2 y_2 + x_3^2 y_3 &= x_4^2 y_4 + x_5^2 y_5 + x_6^2 y_6 \\ x_1 y_1^2 + x_2 y_2^2 + x_3 y_3^2 &= x_4 y_4^2 + x_5 y_5^2 + x_6 y_6^2 \\ y_1^3 + y_2^3 + y_3^3 &= y_4^3 + y_5^3 + y_6^3 \end{aligned} \tag{5.1}$$

with $x_i, y_i \in [1, P] \cap \mathbb{Z}$. Further, write $T(P)$ for the number of "trivial" solutions, in which $(x_4, x_5, x_6) = \sigma(x_1, x_2, x_3)$ and $(y_4, y_5, y_6) = \sigma(y_1, y_2, y_3)$ for some permutation $\sigma \in S_3$. Then clearly one has

$$T(P) = 6P^6 + O(P^4),$$

and it transpires that almost all the solutions counted by $S(P)$ are of this diagonal type. The following "paucity theorem" provides an upper bound for the number of non-diagonal solutions.

Theorem 5.1. *For every $\varepsilon > 0$, one has*

$$S(P) - T(P) \ll_\varepsilon P^{6-\frac{11}{192}+\varepsilon},$$

and hence in particular $S(P) \sim 6P^6$.

Note that the second assertion of the theorem is immediate from the first in view of the above discussion. Section 5.2 is devoted to the proof of Theorem 5.1. In Section 5.3, we use this result, along with the techniques of its proof and the iterative method of Chapter 3, to obtain non-trivial estimates for higher moments. In Section 5.4, we establish Weyl estimates more suitable for small k than those of Theorem 4.1. Finally, in Section 5.5, we apply the circle method as in Sections 4.5 and 4.6 to obtain

Theorem 5.2. *Suppose that $s \geq 58$ and let $L_s(P)$ denote the number of distinct lines ℓ on the hypersurface*

$$c_1 z_1^3 + \cdots + c_s z_s^3 = 0$$

with $h(\ell) \leq P$. Then $L_s(P) \gg P^{2s-12}$ for P sufficiently large.

For comparison, we note that Wooley [68] has demonstrated the existence of rational lines on arbitrary cubic hypersurfaces in at least 37 variables, whereas we require 58 variables in Theorem 5.2. In the additive situation we are considering, the existence of lines follows immediately from the theory of a single additive cubic equation (see R. Baker [6]), provided that $s \geq 14$. Hence the significance of our result lies in the density estimate.

The existence of these "trivial" lines when $s \geq 14$ is in fact key to our analysis, for they give rise to non-singular integer solutions of the system

$$c_1 x_1^{3-j} y_1^j + \cdots + c_s x_s^{3-j} y_s^j = 0 \quad (0 \leq j \leq 3) \tag{5.2}$$

and hence allow us to avoid imposing explicit local solubility hypotheses in Theorem 5.2. Unfortunately, the solutions arising in this way are singular for larger values of k and hence are of no use in the analysis leading to Theorem 8.

Local conditions also may present an obstacle to demonstrating the expected density of higher-dimensional linear spaces on a cubic hypersurface in a reasonable number of variables. While the results of Schmidt [51], [52] could be applied to the analogues of (5.2), the number of variables required may in general be quite large.

The material of this chapter is contained in the author's preliminary manuscript [45].

5.2 The Paucity Problem

Our goal in this section is to establish Theorem 5.1. Before proceeding with the proof, we record for reference some of the key estimates we will use. The first of these is implicit in the work of Hooley [29] on sums of four cubes.

Lemma 5.2.1. *Let n be a non-zero integer, and let $R(P)$ denote the number of integral solutions of the equation*

$$x_1^3 + x_2^3 + x_3^3 + x_4^3 = n$$

with $|x_i| \leq P$. Then one has $R(P) \ll P^{11/6+\varepsilon}$.

Proof. Clearly, we may focus attention on solutions in which at least two of the x_i are non-zero. For any such solution $\mathbf{x}$ counted by $R(P)$, we can find i and j such that x_i and x_j have the same parity and are not both zero. Now if $x_i + x_j = 0$ and x_k and x_l are the remaining two variables, then since $n \neq 0$ we must have $x_k + x_l \neq 0$, and if x_k and x_l do not have the same parity, then one of them has the same parity as x_i and x_j. Thus, after relabeling variables, we may assume that $x_1 \equiv x_2 \pmod 2$ and $x_1 \neq -x_2$. This allows us to write $x_1 = r + s$ and $x_2 = r - s$, where r and s are integers with $r \neq 0$, and hence to consider solutions of the equation

$$2r(r^2 + 3s^2) = n - z^3 - w^3.$$

The argument is now identical to that of Hooley [29], the condition $r \neq 0$ being essential to the consideration of congruences modulo divisors of r. The only change is that the upper bound of $n^{1/3}$ for the moduli of r, z, and w is replaced throughout by P, and the sieving parameter ξ is now chosen to be $P^{1/6}$. $\square$

We also make use of some recent work of Heath-Brown [27] on sums of two cubes.

Lemma 5.2.2. *Let $U(P)$ denote the number of integral solutions of the equation*

$$x_1^3 + x_2^3 = x_3^3 + x_4^3$$

with $|x_i| \leq P$ and $x_1 + x_2 \neq x_3 + x_4$. Then one has $U(P) \ll P^{4/3+\varepsilon}$.

Proof. This is a special case of Heath-Brown [27], Theorem 1. $\square$

We remark that Hooley [30], using the Riemann hypothesis for varieties over finite fields, obtained a result of the above shape with the exponent $4/3$ replaced by $5/3$. Wooley [66] later devised an elementary proof, and his ideas play a key role in Heath-Brown's argument.

Finally, we recall a result on binary quadratic forms dating back to Estermann [22].

Lemma 5.2.3. *Let a, b, and c be non-zero integers, and let $Q(P)$ denote the number of integral solutions of the equation*

$$ax^2 + by^2 = c$$

with $1 \leq x, y \leq P$. Then one has $Q(P) \ll |abcP|^{\varepsilon}$.

Proof. See (for example) Vaughan and Wooley [56], Lemma 3.5. $\square$

We are now ready to embark on the proof of Theorem 5.1. On writing $h = x_1 - x_4$ and $g = y_1 - y_4$ and relabeling variables in (5.1), we see that $S(P)$ is the number of solutions of the system of equations

$$\begin{aligned} h(3x(x+h)+h^2) &= u_1^3 + u_2^3 - u_3^3 - u_4^3 \\ (2hx+h^2)y + g(x+h)^2 &= u_1^2v_1 + u_2^2v_2 - u_3^2v_3 - u_4^2v_4 \\ (2gy+g^2)x + h(y+g)^2 &= u_1v_1^2 + u_2v_2^2 - u_3v_3^2 - u_4v_4^2 \\ g(3y(y+g)+g^2) &= v_1^3 + v_2^3 - v_3^3 - v_4^3 \end{aligned} \tag{5.3}$$

with

$$1 \leq x, y, u_i, v_i \leq P \quad \text{and} \quad |h|, |g| < P. \tag{5.4}$$

We shall estimate $N(P) = S(P) - T(P)$ by dividing into several cases.

(i) Let N_1 denote the number of solutions counted by $N(P)$ for which $h = g = 0$, and consider a solution $x, y, \mathbf{u}, \mathbf{v}$ counted by N_1. Then one has

$$(u_1, u_2, v_1, v_2) \neq (u_3, u_4, v_3, v_4) \quad \text{and} \quad (u_1, u_2, v_1, v_2) \neq (u_4, u_3, v_4, v_3),$$

since otherwise the solution would be counted by $T(P)$. If we have $(u_1, u_2) = (u_3, u_4)$ and $(v_1, v_2) = (v_4, v_3)$, then the second equation in (5.3) implies that either $u_1 = u_2$ or $v_1 = v_2$, whence the number of choices for $\mathbf{u}$ and $\mathbf{v}$ is $O(P^3)$. Trivially, there are $O(P^2)$ choices for x and y, so the total number of solutions is $O(P^5)$, and the same analysis applies if $(u_1, u_2) = (u_4, u_3)$ and $(v_1, v_2) = (v_3, v_4)$. Otherwise, since u_i and v_i are positive, it follows that either $u_1 + u_2 \neq u_3 + u_4$ or $v_1 + v_2 \neq v_3 + v_4$, so Lemma 5.2.2 may be applied to estimate the number of choices for $\mathbf{u}$ or $\mathbf{v}$ (or possibly both). On combining this with Hua's Lemma, one sees that $N_1 \ll P^{16/3+\varepsilon}$.

(ii) Let N_2 denote the number of solutions counted by $N(P)$ for which exactly one of h or g is zero. Suppose first that $h = 0$ and $g \neq 0$. Then by Hua's Lemma one has $O(P^{2+\varepsilon})$ choices for $\mathbf{u}$, and by a trivial estimate there are $O(P^2)$ choices for g and y. Now for fixed non-zero g and y, we may apply Lemma 5.2.1 to deduce that there are

$O(P^{11/6+\varepsilon})$ choices of $\mathbf{v}$ satisfying the fourth equation of (5.3). Finally, since $g \neq 0$, the second equation is a non-trivial polynomial in x and hence determines x to $O(1)$. By following a symmetric argument in the case where $g = 0$ and $h \neq 0$, we find that $N_2 \ll P^{35/6+\varepsilon}$.

(iii) Write $d = (h, g)$, let β be a parameter at our disposal, and let N_3 denote the number of solutions counted by $N(P)$ for which $hg \neq 0$ and $|hg/d| \leq P^{1+\beta}$. In this case, there are $O(P^{1+\beta})$ choices for the integer hg/d, of which d, h/d and g/d are all divisors. Thus by a standard estimate for the divisor function, we see that there are $O(P^{1+\beta+\varepsilon})$ choices for h and g. Trivially, there are $O(P)$ choices for x, and then by Lemma 5.2.1 we have $O(P^{11/6+\varepsilon})$ choices for $\mathbf{u}$. Now by taking a linear combination of the equations (5.3), with respective weights $g^3, -3g^2h, 3gh^2$, and $-h^3$, we find that any solution $x, y, g, h, \mathbf{u}, \mathbf{v}$ satisfies

$$(gu_1 - hv_1)^3 + (gu_2 - hv_2)^3 = (gu_3 - hv_3)^3 + (gu_4 - hv_4)^3, \tag{5.5}$$

and by applying Hua [32], Theorem 4, to the underlying mean value we find that, for fixed h, g, and $\mathbf{u}$, there are $O(P^{2+\varepsilon})$ choices for $\mathbf{v}$. Finally, y is determined to $O(1)$ by a polynomial, whence $N_3 \ll P^{35/6+\beta+\varepsilon}$.

(iv) For $i = 1, \ldots, 4$, write $X_i = gu_i - hv_i$, and let N_4 denote the number of solutions counted by $N(P)$ for which $X_1 + X_2 = X_3 + X_4$ and $hg \neq 0$. The former condition, when combined with (5.5), implies that either $X_1 = X_3$, $X_2 = X_3$, or $X_1 = -X_2$. We may suppose that $X_1 = X_3$ and $X_2 = X_4$, so that

$$g(u_1 - u_3) = h(v_1 - v_3) \quad \text{and} \quad g(u_2 - u_4) = h(v_2 - v_4), \tag{5.6}$$

the argument in the remaining two cases being identical. For convenience we write

$$A = u_1 - u_3, \quad B = u_2 - u_4, \quad C = v_1 - v_3, \quad \text{and} \quad D = v_2 - v_4. \tag{5.7}$$

Since h and g are non-zero, the first equation in (5.3) implies that either A or B is non-zero, and the fourth equation implies that either C or D is non-zero. Suppose that $C \neq 0$ and $D = 0$. We first choose $u_2 = u_4$ and $v_2 = v_4$ in $O(P^2)$ ways, and then by (5.3) we have

$$(x + h)^3 + u_3^3 = x^3 + u_1^3 \quad \text{and} \quad (y + g)^3 + v_3^3 = y^3 + v_1^3.$$

Since solutions with $x = u_3$ and $y = v_3$ are counted by $T(P)$, we may apply Lemma 5.2.2, together with Hua's Lemma, to deduce that there are $O(P^{10/3+\varepsilon})$ choices for $x, y, h, g, u_1, u_3, v_1$, and v_3. The case where $C = 0$ and $D \neq 0$ is identical.

It remains to consider solutions for which both C and D (and hence A and B) are non-zero. We first observe that, after substituting from (5.7) and completing the square, the first and fourth equations in (5.3) become

$$h(3x(x+h)+h^2) - \tfrac{1}{4}(A^3+B^3) = 3A(u_3+\tfrac{1}{2}A)^2 + 3B(u_4+\tfrac{1}{2}B)^2 \tag{5.8}$$

and

$$g(3y(y+g)+g^2) - \tfrac{1}{4}(C^3+D^3) = 3C(v_3+\tfrac{1}{2}C)^2 + 3D(v_4+\tfrac{1}{2}D)^2, \tag{5.9}$$

respectively. In view of (5.8) and (5.9), we further classify solutions according to whether

$$h(3x(x+h)+h^2) - \tfrac{1}{4}(A^3+B^3) = 0 \tag{5.10}$$

or

$$g(3y(y+g)+g^2) - \tfrac{1}{4}(C^3+D^3) = 0. \tag{5.11}$$

If both (5.10) and (5.11) hold, then we start by selecting values for A and B from among $O(P^2)$ possibilities, and (5.10) then determines h and x to $O(P^\varepsilon)$. Trivially, there are $O(P)$ choices for g, and (5.6) then determines C and D to $O(P^\varepsilon)$, whence y is determined to $O(1)$ by (5.11). Finally, u_3 and v_3 may be assigned in $O(P^2)$ ways, and this choice determines $\mathbf{u}$ and $\mathbf{v}$ in light of (5.7), (5.8), and (5.9). Hence there are $O(P^{5+\varepsilon})$ solutions of this type.

If (5.10) holds but (5.11) does not, then we assign A, B, and u_3 in $O(P^3)$ ways, so that $\mathbf{u}$ is determined by (5.8). Then h and x are again determined up to $O(P^\varepsilon)$, and there are $O(P^2)$ choices for y and g. This latter choice determines C and D to $O(P^\varepsilon)$ by (5.6), and we may apply Lemma 5.2.3 to (5.9), regarded as a binary quadratic equation in the variables v_3 and v_4. The case where (5.11) holds but (5.10) does not is exactly symmetric, so we see that there are $O(P^{5+\varepsilon})$ solutions of these two types.

Finally, if neither (5.10) nor (5.11) holds, then we fix h, C, and D in $O(P^3)$ ways, from which g, A, and B are determined to $O(P^\varepsilon)$ by (5.6). There are $O(P^2)$ possibilities for x and y, and Lemma 5.2.3 then shows that $\mathbf{u}$ and $\mathbf{v}$ are determined up to $O(P^\varepsilon)$ by (5.8) and (5.9). Thus we conclude that $N_4 \ll P^{16/3+\varepsilon}$.

(v) Now let γ be a parameter at our disposal, write $M = P^{1+\beta}$, and let N_5 be the number of solutions counted by $N(P)$ in which

$$hg \neq 0, \quad |hg/d| > M, \quad X_1+X_2 \neq X_3+X_4, \tag{5.12}$$

and $d = (h, g) \le P^\gamma$. By symmetry, we may assume that $|h| \ge |g|$, the argument in the other case being identical. Write $h' = h/d$ and $g' = g/d$, so that $(h', g') = 1$. For any given d and $|h'| \ge |g'|$, we divide both sides of (5.5) by d^3 and apply Lemma 5.2.2 to deduce that there are then $O((|h'|P)^{4/3+\varepsilon})$ possible choices for $X_1, \dots, X_4$. With X_i now fixed and $(h', g') = 1$, any two choices for u_i must be congruent modulo h', so one has $O(P/|h'|)$ possibilities for each of $u_1, \dots, u_4$, and this determines $\mathbf{v}$. Since x and y are then determined by polynomials, we find that

$$\begin{aligned} N_5 &\ll \sum_{d \le P^\gamma} \sum_{1 \le g \le P/d} \sum_{h \ge \max(g, M/gd)} (hP)^{4/3+\varepsilon} (P/h)^4 \\ &\ll P^{16/3+\varepsilon} \sum_{d \le P^\gamma} \left(\sum_{g \le (M/d)^{1/2}} \sum_{h \ge M/gd} h^{-8/3} + \sum_{g > (M/d)^{1/2}} \sum_{h \ge g} h^{-8/3} \right) \\ &\ll P^{16/3+\varepsilon} \sum_{d \le P^\gamma} \left(\sum_{g \le (M/d)^{1/2}} (M/gd)^{-5/3} + \sum_{g > (M/d)^{1/2}} g^{-5/3} \right) \\ &\ll P^{16/3+\varepsilon} \sum_{d \le P^\gamma} \left((M/d)^{-5/3} (M/d)^{4/3} + (M/d)^{-1/3} \right) \\ &\ll P^{16/3+\varepsilon} M^{-1/3} \sum_{d \le P^\gamma} d^{1/3} \ll P^{\frac{16}{3} - \frac{1}{3}(1+\beta) + \frac{4}{3}\gamma + \varepsilon}. \end{aligned}$$

(vi) Finally, let N_6 be the number of solutions counted by $N(P)$ with (5.12) and $d > P^\gamma$. In this case we use an affine slicing approach almost exactly as in Wooley [66]. As before, we exploit the symmetry of our system to focus attention on solutions with $|h| \ge |g|$. On recalling (5.5), we have that

$$X_1^3 + X_2^3 = X_3^3 + X_4^3 \quad \text{and} \quad X_1 + X_2 = X_3 + X_4 + H \tag{5.13}$$

for some integer H. For convenience, we write $X_i' = X_i/d$ and $H' = H/d$. For fixed h, g, and $\mathbf{u}$, one has

$$H' = g'(u_1 + u_2 - u_3 - u_4) - h'(v_1 + v_2 - v_3 - v_4),$$

which determines the residue class of H' modulo h'. Furthermore, since $|h'| \ge |g'|$, one has $|H'| \le 4|h'|P$. Now from the equations (5.13), we find that

$$(X_1 + X_2 - X_3)^3 - (X_1^3 + X_2^3 - X_3^3) = (X_4 + H)^3 - X_4^3,$$

which simplifies to

$$3(X_1 - X_3)(X_2 - X_3)(X_1 + X_2) = H(3X_4^2 + 3X_4 H + H^2). \tag{5.14}$$

By (5.12), we have $H \neq 0$, so after dividing both sides of (5.14) by d^3 we see that at least one of $X_1' - X_3'$, $X_2' - X_3'$, or $X_1' + X_2'$ has a divisor $e \gg |H'|^{1/3}$ in common with H'. We suppose that

$$e = (H', X_1' - X_3') \gg |H'|^{1/3}$$

and write $X_1 - X_3 = deY$, the analysis in the other two cases being identical. Hence, for fixed d and e, there are $O(|h'|P/e)$ choices for Y. Now, on substituting $X_4 = X_1 + X_2 - X_3 - H$ and $X_3 = X_1 - deY$ in (5.14), we obtain

$$3deYX_1^2 - 3(deY - H)X_2^2 - 3(deY)^2X_1 - 3(deY - H)^2X_2 = (deY - H)^3 - (deY)^3,$$

and after completing the square this becomes

$$3deY(X_1 - \tfrac{1}{2}deY)^2 - 3(deY - H)(X_2 + \tfrac{1}{2}(deY - H))^2 = \tfrac{1}{4}(deY - H)^3 - \tfrac{1}{4}(deY)^3.$$

Since $H \neq 0$, the quantities deY, $deY - H$, and $(deY - H)^3 - (deY)^3$ are all non-zero, so Lemma 5.2.3 may be applied. Thus, for fixed d, e, H, and Y, the values of X_1 and X_2 are determined up to $O(P^\varepsilon)$, and this fixes X_3 and X_4. For fixed g, h, and $\mathbf{u}$, this determines $\mathbf{v}$, and y is then determined to $O(1)$ by a polynomial. Thus we have

$$N_6 \ll \sum_{d > P^\gamma} \sum_{h', g' \leq P/d} \sum_{x, \mathbf{u}} \sum_{H', e} \frac{h'P^{1+\varepsilon}}{e}.$$

We now divide the sum over H' into dyadic intervals of the form $4h'P/2^{r+1} < |H'| \leq 4h'P/2^r$. For fixed H', a divisor estimate shows that there are $O(P^\varepsilon)$ possible choices for e, and for fixed h and x Lemma 5.2.1 shows that there are $O(P^{11/6+\varepsilon})$ choices for $\mathbf{u}$. Thus on summing trivially over g' and x we find that

$$N_6 \ll \sum_{r=0}^{\infty} \sum_{d > P^\gamma} \frac{P}{d} \sum_{h' \leq P/d} P \cdot P^{11/6+\varepsilon} \cdot \frac{P^{1+\varepsilon}}{2^r} \left(\frac{h'P}{2^{r+1}}\right)^{-1/3} h'P^{1+\varepsilon},$$

on recalling that $e \gg |H'|^{1/3}$ and that the choice of h, g, and $\mathbf{u}$ fixes the residue class of H' modulo h'. Finally, we obtain

$$\begin{aligned} N_6 &\ll P^{11/2+\varepsilon} \sum_{r=0}^{\infty} 2^{-2r/3} \sum_{d > P^\gamma} d^{-1} \sum_{h' \leq P/d} (h')^{2/3} \\ &\ll P^{43/6+\varepsilon} \sum_{d > P^\gamma} d^{-8/3} \ll P^{43/6 - \frac{5}{3}\gamma + \varepsilon}. \end{aligned}$$

We now choose the value of γ so that N_5 and N_6 have the same order of magnitude. Thus we set

$$\frac{16}{3} - \frac{1}{3}(1+\beta) + \frac{4}{3}\gamma = \frac{43}{6} - \frac{5}{3}\gamma,$$

which yields $\gamma = \frac{11}{18} + \frac{1}{9}(1+\beta)$. In view of our bound for N_3, we choose β by setting

$$\frac{35}{6} + \beta = \frac{43}{6} - \frac{5}{3}\left(\frac{11}{18} + \frac{1}{9}(1+\beta)\right),$$

which gives $\beta = \frac{7}{64}$ and $\gamma = \frac{47}{64}$. The result of Theorem 5.1 now follows immediately on assembling the bounds for $N_1, \ldots, N_6$ and noting that $\frac{1}{6} - \frac{7}{64} = \frac{11}{192}$.

5.3 Further Mean Value Estimates

Here we use the result of the previous section to obtain estimates for higher moments, which will be required in our application of the Hardy-Littlewood method in Section 5.5. As usual, the sharpest estimates are obtained by considering solutions in which some of the variables have no large prime factors. Thus when P and R are positive integers, write

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

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

$$f(\boldsymbol{\alpha}; P, R) = \sum_{x,y \in \mathcal{A}(P,R)} e(\alpha_0 x^3 + \alpha_1 x^2 y + \alpha_2 x y^2 + \alpha_3 y^3).$$

It will also be useful to have some variables in a complete interval, so we define

$$F(\boldsymbol{\alpha}; P) = \sum_{1 \le x,y \le P} e(\alpha_0 x^3 + \alpha_1 x^2 y + \alpha_2 x y^2 + \alpha_3 y^3).$$

When there is no danger of confusion, we shall write $f(\boldsymbol{\alpha}) = f(\boldsymbol{\alpha}; P, R)$ and $F(\boldsymbol{\alpha}) = F(\boldsymbol{\alpha}; P)$. Further, let

$$S_{s,r}(P,R) = \int_{\mathbb{T}^4} |F(\boldsymbol{\alpha})|^{2r} |f(\boldsymbol{\alpha})|^{2s} \, d\boldsymbol{\alpha}.$$

We adopt the convention that any statement involving ε and R means that for each $\varepsilon > 0$ there exists $\eta = \eta(\varepsilon) > 0$ such that the assertion holds whenever $R \le P^\eta$. In this section, our implicit constants will depend at most on ε unless otherwise noted. We start with an estimate for $S_{3,2}(P,R)$.

Lemma 5.3.1. *One has*

$$S_{3,2}(P,R) \ll P^{12+\frac{1}{20}+\varepsilon}.$$

Proof. Define the difference operator Δ_1^* by

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

Then by Cauchy's inequality, one has

$$\begin{aligned} S_{3,2}(P,R) &= \int_{\mathbb{T}^4} \left| \sum_{x,h} \sum_{y,g} e\left(\sum_{i=0}^{3} \alpha_i \Delta_1^*(x^{3-i}y^i;h,g) \right) \right|^2 |f(\boldsymbol{\alpha})|^6 \, d\boldsymbol{\alpha} \\ &\le P^2 \sum_{h,g} \int_{\mathbb{T}^4} \left| \sum_{x,y} e\left(\sum_{i=0}^{3} \alpha_i \Delta_1^*(x^{3-i}y^i;h,g) \right) \right|^2 |f(\boldsymbol{\alpha})|^6 \, d\boldsymbol{\alpha}, \end{aligned}$$

and hence

$$S_{3,2}(P,R) \le P^2 U_{3,2}(P,R), \tag{5.15}$$

where $U_{s,2}(P,R)$ denotes the number of solutions of the system

$$\begin{aligned} 3h(x_1^2 - x_2^2 + h(x_1 - x_2)) &= \sum_{i=1}^{s} (u_i^3 - u_{s+i}^3) \\ h(2(x_1y_1 - x_2y_2) + h(y_1 - y_2)) + g(x_1^2 - x_2^2 + 2h(x_1 - x_2)) &= \sum_{i=1}^{s} (u_i^2 v_i - u_{s+i}^2 v_{s+i}) \\ g(2(x_1y_1 - x_2y_2) + g(x_1 - x_2)) + h(y_1^2 - y_2^2 + 2g(y_1 - y_2)) &= \sum_{i=1}^{s} (u_i v_i^2 - u_{s+i} v_{s+i}^2) \\ 3g(y_1^2 - y_2^2 + g(y_1 - y_2)) &= \sum_{i=1}^{s} (v_i^3 - v_{s+i}^3) \end{aligned}$$

with

$$1 \le x_i, y_i \le P, \quad u_i, v_i \in \mathcal{A}(P,R), \quad \text{and} \quad |h|, |g| < P. \tag{5.16}$$

The argument is now very similar to the proof of Theorem 5.1 given in the previous section.

(i) Let U_1 denote the number of solutions counted by $U_{3,2}(P,R)$ for which $h = g = 0$ or $x_1 - x_2 = y_1 - y_2 = 0$. In either case, there are $O(P^4)$ choices for h, g, $\mathbf{x}$, and $\mathbf{y}$, and one sees that the number of choices for $\mathbf{u}$ and $\mathbf{v}$ is then

$$\int_{\mathbb{T}^4} |f(\boldsymbol{\alpha})|^6 \, d\boldsymbol{\alpha} \ll P^6,$$

on recalling Theorem 5.1 and considering the underlying Diophantine equations. Thus we have $U_1 \ll P^{10}$.

(ii) Let U_2 denote the number of solutions counted by $U_{3,2}(P,R)$ for which exactly one of h or g is zero. First suppose that $h=0$, $g\neq 0$, and $y_1\neq y_2$. Then by Vaughan [53], Theorem 4.4, one has $O(P^{13/4+\varepsilon})$ choices for $\mathbf{u}$, and by a trivial estimate there are $O(P^3)$ choices for g and $\mathbf{y}$. Now for fixed g and $\mathbf{y}$, [53] again shows that there are $O(P^{13/4+\varepsilon})$ choices for $\mathbf{v}$. Finally, since $g(y_1-y_2)\neq 0$, x_1 and x_2 are determined to $O(P^{\varepsilon})$ by the second and third equations above. If instead $y_1=y_2$, then one has $O(P^2)$ choices for g and $\mathbf{y}$, but $O(P^{1+\varepsilon})$ choices for $\mathbf{x}$, so we get the same estimate in either case. The situation when $g=0$ and $h\neq 0$ is identical. Thus we see that $U_2 \ll P^{19/2+\varepsilon}$.

(iii) Let U_3 denote the number of solutions counted by $U_{3,2}(P,R)$ for which $hg\neq 0$ and exactly one of x_1-x_2 or y_1-y_2 is zero. If $x_1\neq x_2$, and $y_1=y_2$, then there are $O(P^{13/4+\varepsilon})$ choices for $\mathbf{v}$ and $O(P^2)$ choices for h and g. Now, as in the previous section, we find that

$$\begin{aligned}(gu_1-hv_1)^3+(gu_2-hv_2)^3+(gu_3-hv_3)^3\\ =(gu_4-hv_4)^3+(gu_5-hv_5)^3+(gu_6-hv_6)^3\end{aligned} \tag{5.17}$$

so by Hua [32], Theorem 4, there are $O(P^{7/2+\varepsilon})$ choices for $\mathbf{u}$, and then $\mathbf{x}$ and $\mathbf{y}$ are determined to $O(P^{1+\varepsilon})$. The other case is identical, and thus $U_3 \ll P^{39/4+\varepsilon}$.

(iv) Write $d=(h,g)$, let β be a parameter at our disposal, and let U_4 denote the number of solutions counted by $U_{3,2}(P,R)$ for which

$$0<|hg/d|\leq P^{5/4+\beta},\quad x_1\neq x_2,\quad \text{and}\quad y_1\neq y_2.$$

In this case, there are $O(P^{5/4+\beta})$ choices for the integer hg/d, whence by a divisor estimate there are $O(P^{5/4+\beta+\varepsilon})$ choices for h and g. Trivially, there are $O(P^2)$ choices for $\mathbf{x}$, and then we have $O(P^{13/4+\varepsilon})$ choices for $\mathbf{u}$. Thus by applying Hua [32] to (5.17), we see that there are $O(P^{7/2+\varepsilon})$ choices for $\mathbf{v}$, and then $\mathbf{y}$ is determined to $O(P^{\varepsilon})$. Hence $U_4 \ll P^{10+\beta+\varepsilon}$.

(v) Now let γ be a parameter at our disposal, write $M=P^{5/4+\beta}$, and let U_5 be the number of solutions counted by $U_{3,2}(P,R)$ in which

$$hg\neq 0,\quad |hg/d|>M,\quad x_1\neq x_2,\quad y_1\neq y_2, \tag{5.18}$$

and $d=(h,g)\leq P^{\gamma}$. As before, we assume that $|h|\geq|g|$ and write $h'=h/d$ and $g'=g/d$, so that $(h',g')=1$. For any given d and $|h'|\geq|g'|$, we divide both sides of (5.17) by d^3; then by Hua's Lemma there are $O((|h'|P)^{7/2+\varepsilon})$ possible choices for

$X_1, \ldots, X_6$, where $X_i = gu_i - hv_i$. With X_i now fixed and $(h', g') = 1$, any two choices for u_i must be congruent modulo h', so one has $O(P/|h'|)$ possibilities for each of $u_1, \ldots, u_6$, and this determines $\mathbf{v}$. Since $\mathbf{x}$ and $\mathbf{y}$ are then determined to $O(P^\varepsilon)$, we find that

$$\begin{aligned}
U_5 &\ll \sum_{d \le P^\gamma} \sum_{1 \le g \le P/d} \sum_{h \ge \max(g, M/gd)} (hP)^{7/2+\varepsilon} (P/h)^6 \\
&\ll P^{19/2+\varepsilon} \sum_{d \le P^\gamma} \left(\sum_{g \le (M/d)^{1/2}} \sum_{h \ge M/gd} h^{-5/2} + \sum_{g > (M/d)^{1/2}} \sum_{h \ge g} h^{-5/2} \right) \\
&\ll P^{19/2+\varepsilon} \sum_{d \le P^\gamma} \left(\sum_{g \le (M/d)^{1/2}} (M/gd)^{-3/2} + \sum_{g > (M/d)^{1/2}} g^{-3/2} \right) \\
&\ll P^{19/2+\varepsilon} \sum_{d \le P^\gamma} \left((M/d)^{-3/2} (M/d)^{5/4} + (M/d)^{-1/4} \right) \\
&\ll P^{19/2+\varepsilon} M^{-1/4} \sum_{d \le P^\gamma} d^{1/4} \ll P^{\frac{19}{2} - \frac{1}{4}(\frac{5}{4}+\beta) + \frac{5}{4}\gamma + \varepsilon}.
\end{aligned}$$

(vi) Finally, let U_6 be the number of solutions counted by $U_{3,2}(P, R)$ with (5.18) and $d > P^\gamma$. Then for fixed d, there are $(P/d)^2$ choices for h and g and P^2 choices for $\mathbf{x}$. Now on recalling (5.17), we see that there are $O(P^{27/4+\varepsilon})$ choices for $\mathbf{u}$ and $\mathbf{v}$, whence $\mathbf{y}$ is determined to $O(P^\varepsilon)$. Thus

$$U_6 \ll P^{43/4+\varepsilon} \sum_{d > P^\gamma} d^{-2} \ll P^{43/4-\gamma+\varepsilon}.$$

To optimize the results of (v) and (vi), we set

$$9 + \frac{3}{16} - \frac{1}{4}\beta + \frac{5}{4}\gamma = 10 + \frac{3}{4} - \gamma,$$

which gives $\gamma = \frac{25}{36} + \frac{1}{9}\beta$. Now on recalling (iv), we choose β so that

$$10 + \beta = \frac{43}{4} - \frac{25}{36} - \frac{1}{9}\beta,$$

which gives $\beta = \frac{1}{20}$. Hence $U_{3,2}(P, R) \ll P^{10+\frac{1}{20}}$, and the lemma follows on recalling (5.15). $\square$

Before proceeding, we record an easy consequence of Lemma 5.3.1.

Lemma 5.3.2. *One has*

$$S_{2,2}(P, R) \ll P^{9+\frac{1}{40}+\varepsilon}.$$

Proof. By Theorem 5.1, we have

$$\int_{\mathbb{T}^4} |F(\boldsymbol{\alpha})|^6 d\boldsymbol{\alpha} \ll P^6.$$

Thus by applying the Cauchy-Schwarz inequality and Lemma 5.3.1 we obtain

$$S_{2,2}(P,R) \ll \left(\int_{\mathbb{T}^4} |F(\boldsymbol{\alpha})|^6 d\boldsymbol{\alpha} \right)^{1/2} \left(\int_{\mathbb{T}^4} |F(\boldsymbol{\alpha})|^2 |f(\boldsymbol{\alpha})|^8 d\boldsymbol{\alpha} \right)^{1/2} \ll P^{9+\frac{1}{40}+\varepsilon}$$

on considering the underlying Diophantine equations. □

We now proceed to estimate some higher moments.

Lemma 5.3.3. *One has*

$$S_{4,2}(P,R) \ll P^{15+\frac{1}{3}+\varepsilon}.$$

Proof. By Cauchy's inequality, we have $S_{4,2}(P,R) \leq P^2 U_{4,2}(P,R)$, and the estimation of $U_{4,2}(P,R)$ proceeds almost exactly as in the proof of Lemma 5.3.1. The only modifications are that we use Lemma 5.3.2 in place of Theorem 5.1 in the analysis of case (i) and we replace the 6th moment estimates of $P^{13/4+\varepsilon}$ and $P^{7/2+\varepsilon}$ by Hua's 8th moment estimate of $P^{5+\varepsilon}$. Taking $M = P^{4/3}$ and $\gamma = 2/3$ produces identical bounds for the final three cases and hence gives an optimal result. □

We remark that the argument of the preceding proof in fact shows that one may replace $S_{4,2}(P,R)$ by $S_{2,4}(P,R)$ in the statement of Lemma 5.3.3. Further, we note that tiny improvements in the exponents of the preceding lemmata may be achieved by using the results of Wooley [64], [71] in place of Vaughan [53], but such improvements do not have significant consequences in the present application.

For higher moments, we apply the (single) efficient differencing procedure of Chapter 3. Although our methods always allow us to take a few variables ranging over a complete interval, we will often simplify by stating results for mean values in which all the variables are smooth. Thus we adopt the notation of writing $S_s(P,R)$ for $S_{s,0}(P,R)$. Further, we say that Δ_s is an admissible exponent if one has $S_s(P,R) \ll P^{\lambda_s+\varepsilon}$, where $\lambda_s = 4s - 12 + \Delta_s$, and in this situation we call λ_s a permissible exponent. To this point we have obtained the admissible exponents

$$\Delta_3 = 6, \quad \Delta_4 = 5\tfrac{1}{40}, \quad \Delta_5 = 4\tfrac{1}{20}, \quad \text{and} \quad \Delta_6 = 3\tfrac{1}{3}. \tag{5.19}$$

The above method of generating admissible exponents becomes noticeably less effective when $s > 6$, since the maximum savings of P^3 in estimating the number of solutions of

$$u_1^3 + \cdots + u_t^3 = u_{t+1}^3 + \cdots + u_{2t}^3$$

is already achieved when $t = 4$. Since the results of Section 3.5 are directly applicable only for $s \geq 11$, we will need the following lemma to work out admissible exponents when s lies in the intermediate range.

Lemma 5.3.4. *One has*

$$S_{s+2}(P,R) \ll P^5 S_s(P,R) + P^{19/6+\varepsilon} S_{s-1,2}(P,R) + P^{\frac{2}{3}s+6+\varepsilon} S_s(P^{5/6}, R).$$

Proof. This follows on using Lemma 3.3.2 in the initial stages of the argument of Lemma 3.5.1. $\square$

We now apply Lemma 5.3.4 repeatedly to obtain admissible exponents Δ_s for $7 \leq s \leq 12$. First of all, by using Lemma 5.3.3 and making a trivial estimate, we see that

$$S_{5,2}(P,R) \ll P^{19+\frac{1}{3}+\varepsilon},$$

and using this together with Lemma 5.3.3 in Lemma 5.3.4 gives

$$S_8(P,R) \ll P^{22+\frac{7}{9}+\varepsilon}. \tag{5.20}$$

Now using the Cauchy-Schwarz inequality to interpolate between $S_{4,2}$ and S_8, we obtain

$$S_7(P,R) \ll (S_{4,2}(P,R))^{1/2}(S_8(P,R))^{1/2} \ll P^{19+\frac{1}{18}+\varepsilon}. \tag{5.21}$$

Putting (5.20) and (5.21) into Lemma 5.3.4 now yields

$$S_9(P,R) \ll P^{26+\frac{59}{108}+\varepsilon},$$

and this is used along with (5.20) to obtain

$$S_{10}(P,R) \ll P^{30+\frac{17}{54}+\varepsilon}.$$

Continuing the iteration, we find that

$$S_{11}(P,R) \ll P^{34+\frac{79}{648}+\varepsilon} \quad \text{and} \quad S_{12}(P,R) \ll P^{37+\frac{301}{324}+\varepsilon}.$$

Thus we have the admissible exponents

$$\begin{gathered} \Delta_7 = 3\tfrac{1}{18}, \quad \Delta_8 = 2\tfrac{7}{9}, \quad \Delta_9 = 2\tfrac{59}{108}, \\ \Delta_{10} = 2\tfrac{17}{54}, \quad \Delta_{11} = 2\tfrac{79}{648}, \quad \Delta_{12} = \tfrac{625}{324}. \end{gathered} \tag{5.22}$$

Further admissible exponents can now be read off from Lemma 3.5.1. Namely, if $s \geq 11$ and Δ_s is admissible then the exponent $\Delta_{s+2t} = \Delta_s(5/6)^t$ is also admissible.

5.4 Weyl Differencing

Here we obtain estimates for the modulus of the exponential sum $F(\boldsymbol{\alpha})$ when at least one of the α_j is badly approximated by rationals. In Section 4.2, estimates of this type were obtained for $f(\boldsymbol{\alpha})$ by using the large sieve to relate the modulus of the sum to known mean values. This treatment allowed us to obtain bounds of the form $P^{2-\sigma(k)+\varepsilon}$, where $\sigma(k)^{-1} \asymp k^3 \log k$, and for large k this is substantially better than the exponential decay that results from Weyl differencing. For $k = 3$, however, we are much better off applying a two-fold Weyl differencing procedure. For purposes of application, it is useful to consider the slightly more general exponential sum

$$F(\boldsymbol{\alpha}; P, Q) = \sum_{1\le x\le P} \sum_{1\le y\le Q} e(\alpha_0 x^3 + \alpha_1 x^2 y + \alpha_2 x y^2 + \alpha_3 y^3).$$

Lemma 5.4.1. *Suppose that $Q \asymp P$ and that for some j there exist $q_j \in \mathbb{N}$ and $a_j \in \mathbb{Z}$ with*

$$(q_j, a_j) = 1 \quad \textit{and} \quad |q_j\alpha_j - a_j| \le q_j^{-1}. \tag{5.23}$$

Then one has

$$|F(\boldsymbol{\alpha}; P, Q)| \ll P^{2+\varepsilon}(q_j^{-1} + P^{-1} + q_j P^{-3})^{1/4}.$$

Proof. First suppose that (5.23) holds with $j = 0$. Then by Weyl's inequality (see for instance Lemma 2.4 of Vaughan [55]) one has

$$\begin{aligned} |F(\boldsymbol{\alpha}; P, Q)| &\le \sum_{y\le Q} \left| \sum_{x\le P} e(\alpha_0 x^3 + \alpha_1 x^2 y + \alpha_2 x y^2) \right| \\ &\ll QP^{1+\varepsilon}(q_0^{-1} + P^{-1} + q_0 P^{-3})^{1/4}, \end{aligned}$$

and the result follows. Note that if instead (5.23) holds with $j = 3$, then we obtain the same conclusion simply by interchanging the roles of x and y in the above argument.

Now suppose that (5.23) holds with $j = 1$. Then by Cauchy's inequality we have

$$\begin{aligned} |F(\boldsymbol{\alpha})|^2 &\le P \sum_{y\le Q} \left| \sum_{x\le P} e(\alpha_0 x^3 + \alpha_1 x^2 y + \alpha_2 x y^2) \right|^2 \\ &\ll P \sum_{y\le Q} \left(P + \sum_{x,h} e(\alpha_0(3x^2 h + 3xh^2 + h^3) + \alpha_1 y(2xh + h^2) + \alpha_2 y^2 h) \right) \\ &\ll P^3 + P \sum_{y\le Q} \sum_{x,h} e(\alpha_0(3x^2 h + 3xh^2 + h^3) + \alpha_1 y(2xh + h^2) + \alpha_2 y^2 h), \end{aligned}$$

where the second sum is over x and $h \neq 0$ with $1 \leq x \leq P$ and $1-x \leq h \leq P-x$, and where we have abbreviated $F(\boldsymbol{\alpha}; P, Q)$ by $F(\boldsymbol{\alpha})$. Then on using Cauchy's inequality again we obtain

$$\begin{aligned}
|F(\boldsymbol{\alpha})|^4 &\ll P^6 + P^4 \sum_{x,h} \left| \sum_{y \leq Q} e(\alpha_0(3x^2h + 3xh^2 + h^3) + \alpha_1 y(2xh + h^2) + \alpha_2 y^2 h) \right|^2 \\
&\ll P^6 + P^4 \sum_{x,h} \left(P + \sum_{y,g} e(\alpha_1 g(2xh + h^2) + \alpha_2 h(2yg + g^2)) \right) \\
&\ll P^7 + P^4 \sum_{x,h,y,g} e(\alpha_1 g(2xh + h^2) + \alpha_2 h(2yg + g^2)) \\
&\ll P^7 + P^4 \sum_{1 \leq |h|, |g| \leq P} \sum_{y \in I(Q,g)} \left| \sum_{x \in I(P,h)} e(2\alpha_1 ghx) \right|,
\end{aligned}$$

where $I(P,h) = [1,P] \cap [1-h, P-h]$. Thus on summing the geometric progression, recalling a standard divisor function estimate, and using Lemma 2.2 of [55], we find that

$$\begin{aligned}
|F(\boldsymbol{\alpha}; P, Q)|^4 &\ll P^7 + QP^4 \sum_{1 \leq |h|, |g| \leq P} \min(P, ||2\alpha_1 gh||^{-1}) \\
&\ll P^7 + QP^{4+\varepsilon} \sum_{n \leq 2P^2} \min(P, ||\alpha_1 n||^{-1}) \\
&\ll P^7 + QP^{7+\varepsilon}(q_1^{-1} + P^{-1} + q_1 P^{-3}),
\end{aligned}$$

whence

$$|F(\boldsymbol{\alpha}; P, Q)| \ll P^{2+\varepsilon}(q_1^{-1} + P^{-1} + q_1 P^{-3})^{1/4}.$$

Again, the same conclusion follows when (5.23) holds with $j = 2$ by repeating the argument with the roles of x and y reversed. □

Next we record a consequence of the above lemma, which will be useful in our application of the circle method.

Lemma 5.4.2. *Let $\alpha_0, \ldots, \alpha_3$ be real numbers with the property that whenever there exist $q \in \mathbb{N}$ and $a_0, \ldots, a_3 \in \mathbb{Z}$ with $(q, a_0, \ldots, a_3) = 1$ and $|q\alpha_j - a_j| \leq P^{\delta - 3}$ one has $q > P^{\delta}$. Then whenever $Q \asymp P$ one has*

$$|F(\boldsymbol{\alpha}; P, Q)| \ll P^{2 - \delta/16 + \varepsilon}.$$

Proof. Let $\alpha_0, \ldots, \alpha_3$ be as in the statement of the lemma, and write $\nu = \delta/4$. For each j, Dirichlet's Theorem allows us to find $q_j \in \mathbb{N}$ and $b_j \in \mathbb{Z}$ with $(q_j, b_j) = 1$ such that $|q_j\alpha_j - b_j| \le P^{\nu-3}$ and $q_j \le P^{3-\nu}$. If $q_j > P^\nu$ for some j, then the conclusion follows from Lemma 5.4.1. Otherwise, put $q = [q_0, \ldots, q_3]$ and $a_j = b_j q/q_j$. Then we have $(q, a_0, \ldots, a_3) = 1$ and $q \le q_j P^{3\nu}$ for each j and hence

$$|\alpha_j - a_j/q| \le q_j^{-1} P^{\nu-3} \le q^{-1} P^{\delta-3} \quad (0 \le j \le 3)$$

and $q \le P^\delta$, contradicting the hypothesis of the lemma. □

5.5 The Circle Method

Now we are in a position to prove Theorem 5.2 by applying the circle method along the lines of Section 4.6. The following lemma provides us with non-singular local solutions to (5.2).

Lemma 5.5.1. *If $s \ge 14$ and $c_1, \ldots, c_s$ are non-zero integers, then the system* (5.2) *has a non-singular real solution and a non-singular p-adic solution for all primes p.*

Proof. After setting

$$y_1 = \cdots = y_7 = 0, \quad x_8 = \cdots = x_{14} = 0, \quad \text{and} \quad x_i = y_i = 0 \quad (i > 14),$$

the system (5.2) reduces to

$$c_1 x_1^3 + \cdots + c_7 x_7^3 = 0 \quad \text{and} \quad c_8 y_8^3 + \cdots + c_{14} y_{14}^3 = 0.$$

By Baker [6], each of these equations has a non-trivial integral solution; suppose that $\mathbf{x}$ and $\mathbf{y}$ are solutions with x_I and y_J non-zero. Then on writing

$$\psi_j(\mathbf{x}, \mathbf{y}) = c_1 x_1^{3-j} y_1^j + \cdots + c_s x_s^{3-j} y_s^j \quad (0 \le j \le 3),$$

we have

$$\det \left(\frac{\partial \psi_j}{\partial x_I}, \frac{\partial \psi_j}{\partial y_I}, \frac{\partial \psi_j}{\partial x_J}, \frac{\partial \psi_j}{\partial y_J} \right)_{0 \le j \le 3} = (3 c_I c_J)^2 x_I^4 y_J^4 \ne 0.$$

Thus $(\mathbf{x}, \mathbf{y})$ is a non-singular integer solution of (5.2), so it is non-singular in each local field as well. □

By Lemmata 4.6.2 and 5.5.1, we may assume that the system (5.2) has a non-singular real solution $(\boldsymbol{\eta}, \boldsymbol{\xi})$ with $0 < |\eta_i|, |\xi_i| < \frac{1}{2}$ for $i = 1, \ldots, s$. For each i, we write

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

and

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

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

$$F_i(\boldsymbol{\alpha}) = \sum_{\eta_i^- P < x \le \eta_i^+ P} \sum_{\xi_i^- P < y \le \xi_i^+ P} e(c_i(\alpha_0 x^3 + \alpha_1 x^2 y + \alpha_2 x y^2 + \alpha_3 y^3))$$

and

$$f_i(\boldsymbol{\alpha}) = \sum_{\substack{\eta_i^- P < x \le \eta_i^+ P \\ |x| \in \mathcal{A}(P,R)}} \sum_{\substack{\xi_i^- P < y \le \xi_i^+ P \\ |y| \in \mathcal{A}(P,R)}} e(c_i(\alpha_0 x^3 + \alpha_1 x^2 y + \alpha_2 x y^2 + \alpha_3 y^3)).$$

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

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

Finally, let

$$R_s(P) = \int_{\mathbb{T}^4} \mathcal{F}(\boldsymbol{\alpha}) \mathcal{G}(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha},$$

and observe that $R_s(P)$ is a lower bound for the number of integral solutions of (5.2) lying in the box $[-P, P]^{2s}$.

We dissect $\mathbb{T}^4$ into major and minor arcs as follows. Let $c \in \mathbb{N}$ and $\delta \in (0, 1]$ be parameters at our disposal, and define

$$\mathfrak{M} = \bigcup_{\substack{1 \le a_0, \ldots, a_3 \le q \le cP^\delta \\ (q, a_0, \ldots, a_3) = 1}} \mathfrak{M}(q, \mathbf{a}),$$

where

$$\mathfrak{M}(q, \mathbf{a}) = \{\boldsymbol{\alpha} \in \mathbb{T}^4 : |q\alpha_i - a_i| \le P^{\delta - 3} \ (0 \le i \le 3)\},$$

and put $\mathfrak{m} = \mathbb{T}^4 \setminus \mathfrak{M}$. It is easily seen that the $\mathfrak{M}(q, \mathbf{a})$ are disjoint.

As in Chapter 4, we further define the pruned major arcs by

$$\mathfrak{N} = \bigcup_{\substack{1 \le a_0, \ldots, a_3 \le q \le W \\ (q, a_0, \ldots, a_3) = 1}} \mathfrak{N}(q, \mathbf{a}),$$

where W is a parameter at our disposal and

$$\mathfrak{N}(q, \mathbf{a}) = \{\boldsymbol{\alpha} \in \mathbb{T}^4 : |\alpha_i - a_i/q| \le W P^{-3} \ (0 \le i \le 3)\}.$$

The following pruning lemma is essentially Lemma 9.2 of Wooley [61].

Lemma 5.5.2. *If* $t \geq \max(16, \frac{5\delta}{1-\delta})$, *then one has*

$$\int_{\mathfrak{M}} |F_i(\boldsymbol{\alpha})|^t d\boldsymbol{\alpha} \ll P^{2t-12}$$

and

$$\int_{\mathfrak{M}\setminus\mathfrak{N}} |F_i(\boldsymbol{\alpha})|^t d\boldsymbol{\alpha} \ll W^{-\sigma} P^{2t-12},$$

for some $\sigma > 0$.

Proof. We use Lemma 4.4.1 in the argument of Lemma 4.6.7. □

We are now finally in a position to derive a mean value estimate in which we obtain the full savings of P^{12} over the trivial bound.

Lemma 5.5.3. *One has*

$$\int_{\mathbb{T}^4} |F(\boldsymbol{\alpha})|^{17} |f(\boldsymbol{\alpha})|^{44} d\boldsymbol{\alpha} \ll P^{110}.$$

Proof. Dissect into major and minor arcs by taking $\delta = 3/4$ and $c = 1$ in the above definitions. Then by Lemma 5.4.2 we have

$$\sup_{\boldsymbol{\alpha}\in\mathfrak{m}} |F(\boldsymbol{\alpha})| \ll P^{2-\frac{3}{64}+\varepsilon},$$

and by Lemma 3.5.1, together with (5.22), we see that the exponent

$$\Delta_{22} = \frac{625}{324}\left(\frac{5}{6}\right)^5 < \frac{51}{64}$$

is admissible. Thus by Lemma 5.5.2 and a change of variables we have

$$\begin{aligned}\int_{\mathbb{T}^4} |F(\boldsymbol{\alpha})|^{17} |f(\boldsymbol{\alpha})|^{44} d\boldsymbol{\alpha} &\ll P^{34-\frac{51}{64}+\varepsilon} \int_{\mathbb{T}^4} |f(\boldsymbol{\alpha})|^{44} d\boldsymbol{\alpha} + P^{88} \int_{\mathfrak{M}} |F(\boldsymbol{\alpha})|^{17} d\boldsymbol{\alpha} \\ &\ll P^{110-\frac{51}{64}+\Delta_{22}+\varepsilon} + P^{110},\end{aligned}$$

and the result follows. □

We are now ready to derive the lower bound for $R_s(P)$. Take $t = 24$, $u = 16$, and $v \geq 2$, for a total of $s \geq 58$ variables, and further let $\delta = 9/10$ and $c = \max |c_i|$.

If $\boldsymbol{\alpha} \in \mathfrak{m}$, then by the argument of Lemma 4.6.3 one sees that the hypotheses of Lemma 5.4.2 hold with $\boldsymbol{\alpha}$ replaced by $c_i\boldsymbol{\alpha}$ and hence

$$\sup_{\boldsymbol{\alpha}\in\mathfrak{m}} |F_i(\boldsymbol{\alpha})| \ll P^{2-\delta/16+\varepsilon}.$$

Moreover, the exponent

$$\Delta_{16} = \frac{625}{324}\left(\frac{5}{6}\right)^2 < 1.34$$

is admissible, and it is clear by considering the underlying Diophantine equations that all the mean value estimates from Section 5.3 hold with $f(\boldsymbol{\alpha})$ replaced by $f_i(\boldsymbol{\alpha})$. Thus by Hölder's inequality we have

$$\begin{aligned}\int_{\mathfrak{m}} |\mathcal{F}(\boldsymbol{\alpha})\mathcal{G}(\boldsymbol{\alpha})|\, d\boldsymbol{\alpha} &\ll P^{48+2v-\frac{3}{2}\delta+\varepsilon} \prod_{i=25}^{56} \left(\int_{\mathbb{T}^4} |f_i(\boldsymbol{\alpha})|^{32} d\boldsymbol{\alpha}\right)^{1/32} \\ &\ll P^{2s-12-\frac{27}{20}+\Delta_{16}+\varepsilon} \ll P^{2s-12-\tau}\end{aligned}$$

for some $\tau > 0$. Furthermore, after applying Hölder's inequality and making a change of variables, we have by Lemmata 5.5.2 and 5.5.3 that

$$\begin{aligned}\int_{\mathfrak{M}\setminus\mathfrak{N}} |\mathcal{F}(\boldsymbol{\alpha})\mathcal{G}(\boldsymbol{\alpha})|\, d\boldsymbol{\alpha} &\ll P^{2v} \left(\int_{\mathfrak{M}\setminus\mathfrak{N}} |F(\boldsymbol{\alpha})|^{48} d\boldsymbol{\alpha}\right)^{1/2} \left(\int_{\mathbb{T}^4} |f(\boldsymbol{\alpha})|^{64} d\boldsymbol{\alpha}\right)^{1/2} \\ &\ll P^{2s-12} W^{-\sigma/2},\end{aligned}$$

and so it suffices to deal with the pruned major arc $\mathfrak{N}$. But it follows immediately from the analysis of Section 4.6 that

$$\int_{\mathfrak{N}} \mathcal{F}(\boldsymbol{\alpha})\mathcal{G}(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} \gg P^{2s-12},$$

provided that W is taken to be a suitably small power of $\log P$. Finally, on recalling the notation of Section 4.6 we have by the argument of Lemma 4.6.12 that

$$R_s(P,d) \ll \frac{P^{2v}}{d^v} \int_{\mathbb{T}^4} |\mathcal{F}(\boldsymbol{\alpha})| \left(\prod_{i=25}^{56} |f_i(\boldsymbol{\alpha})|\right) d\boldsymbol{\alpha} \ll \frac{P^{2s-12}}{d^2},$$

since $v \geq 2$. Thus on following through the corresponding argument in the proof of Theorem 4.5, we find that $L_s(P) \gg P^{2s-12}$, and this completes the proof of Theorem 5.2.

We remark that essentially the same analysis, modified along the lines of Section 4.5, may be applied to prove Theorem 7. In that argument, we may clearly take $v = 0$ and it follows that $G_1^*(3) \leq 56$.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] G. I. Arkhipov, *On the Hilbert-Kamke problem*, Math. USSR Izvestiya **24 (1)** (1985), 1–47.

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

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

[4] R. C. Baker, *Cubic Diophantine inequalities*, Mathematika **29** (1982), 83–92.

[5] ———, *Diophantine inequalities*, Clarendon Press, Oxford, 1986.

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

[7] R. C. Baker, J. Brüdern, and T. D. Wooley, *Cubic Diophantine inequalities*, Mathematika **42** (1995), 264–277.

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

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

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

[11] J. Brüdern, *Additive Diophantine inequalities with mixed powers I*, Mathematika **34** (1987), 124–130.

[12] ———, *Cubic Diophantine inequalities*, Mathematika **35** (1988), 51–58.

[13] ———, *Cubic Diophantine inequalities II*, J. London Math. Soc. **53** (1996), 1–18.

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

[15] R. J. Cook, *Simultaneous quadratic inequalities*, Acta Arith. **25** (1974), 337–346.

[16] H. Davenport, *On Waring's problem for fourth powers*, Ann. Math. **40** (1939), 731–747.

[17] ———, *On indefinite quadratic forms in many variables*, Mathematika **3** (1956), 81–101.

[18] ______, *Analytic methods for Diophantine equations and inequalities*, Ann Arbor Publishers, 1962.

[19] ______, *Multiplicative number theory*, Springer-Verlag, New York, 1980.

[20] H. Davenport and H. Heilbronn, *On indefinite quadratic forms in five variables*, J. London Math. Soc. **21** (1946), 185–193.

[21] L. E. Dickson, *All integers except* 23 *and* 239 *are sums of eight cubes*, Bull. Am. Math. Soc. **45** (1939), 588–591.

[22] T. Estermann, *Einige sätze über quadratfreie zahlen*, Math. Ann. **105** (1931), 653–662.

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

[24] G. H. Hardy and J. E. Littlewood, *Some problems of partitio numerorum I: A new solution of Waring's problem*, Göttingen Nachrichten (1920), 33–54.

[25] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 5th ed., Clarendon Press, Oxford, 1979.

[26] G. Harman, *Trigonometric sums over primes I*, Mathematika **28** (1981), 249–254.

[27] D. R. Heath-Brown, *The density of rational points on cubic hypersurfaces*, Acta Arith. **79** (1997), 17–30.

[28] D. Hilbert, *Beweis für die Darstellbarkeit der ganzen zahlen durch eine feste Anzahl* n^{ter} *Potenzen (Waringsches Problem)*, Mat. Annalen **67** (1909), 281–300.

[29] C. Hooley, *On the representations of a number as the sum of four cubes*, Proc. London Math. Soc. **(3) 36** (1978), 117–140.

[30] ______, *On the numbers that are representable as the sum of two cubes*, J. reine angew. Math. **314** (1980), 146–173.

[31] L.-K. Hua, *On Waring's problem*, Q. J. Math. **9** (1938), 199–202.

[32] ______, *Additive theory of prime numbers*, A.M.S., Providence, Rhode Island, 1965.

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

[34] A. Kempner, *Bemerkungen zum Waringschen Problem*, Mat. Annalen **72** (1912), 387–399.

[35] E. Landau, *Über eine Anwendung der Primzahltheorie auf das Waringsche Problem in der elementaren Zahlentheorie*, Mat. Annalen **66** (1909), 102–106.

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

[37] Yu V. Linnik, *On the representation of large numbers as sums of seven cubes*, Mat. Sbornik NS **12** (1943), 218–224.

[38] J. H. Loxton and R. C. Vaughan, *The estimation of complete exponential sums*, Canad. Math. Bull. **28 (4)** (1985), 440–454.

[39] G. A. Margulis, *Indefinite quadratic forms and unipotent flows on homogeneous spaces*, C. R. Acad. Sci. Paris Ser. I Math. **304** (1987), 249–253.

[40] A. Meyer, Vierteljahrsschrift der Naturforschenden Ges. in Zürich **29** (1884), 209–222.

[41] R. Miranda, *Linear systems of plane curves*, Notices Amer. Math. Soc. **46 (2)** (1999), 192–202.

[42] H. L. Montgomery, *Ten lectures on the interface between analytic number theory and harmonic analysis*, AMS, Providence, 1994.

[43] M. B. Nathanson, *Additive number theory: The classical bases*, Springer-Verlag, New York, 1996.

[44] I. Niven, H. S. Zuckerman, and H. L. Montgomery, *An introduction to the theory of numbers*, 5th ed., Wiley, 1991.

[45] S. T. Parsell, *The density of rational lines on cubic hypersurfaces*, (in preparation).

[46] ———, *Multiple exponential sums over smooth numbers*, (submitted).

[47] ———, *On simultaneous diagonal inequalities*, J. London Math. Soc. (to appear).

[48] J. Pitman, *Pairs of diagonal inequalities*, Recent Progress in Analytic Number Theory (H. Halberstam and C. Hooley, eds.), vol. 2, Academic Press, London, 1981, pp. 183–215.

[49] J. Pitman and D. Ridout, *Diagonal cubic equations and inequalities*, Proc. Roy. Soc. London Ser. A **297** (1967), 476–502.

[50] K. F. Roth, *Rational approximations to algebraic numbers*, Mathematika **2** (1955), 1–20.

[51] W. M. Schmidt, *On cubic polynomials III. Systems of p-adic equations*, Monatsh. Math. **93** (1982), 211–223.

[52] ———, *On cubic polynomials IV. Systems of rational equations*, Monatsh. Math. **93** (1982), 329–348.

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

[54] ———, *The use in additive number theory of numbers without large prime factors*, Phil. Trans. R. Soc. Lond. A **345** (1993), 363–376.

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

[56] R. C. Vaughan and T. D. Wooley, *Further improvements in Waring's problem*, Acta Math. **174** (1995), 147–240.

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

[58] ______, *On an upper bound for $G(n)$*, Izv. Akad. Nauk. SSSR **23** (1959), 637–642.

[59] H. Weyl, *Über die Gleichverteilung von Zahlen mod Eins*, Math. Ann. **77** (1916), 313–352.

[60] A. Wieferich, *Beweis des Satzes daß sich eine jede ganze Zahl als Summe von höchstens neun positiven Kuben darstellen läßt*, Mat. Annalen **66** (1909), 95–101.

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

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

[63] ______, *On Vinogradov's mean value theorem*, Mathematika **39** (1992), 379–399.

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

[65] ______, *New estimates for smooth Weyl sums*, J. London Math. Soc. (2) **51** (1995), 1–13.

[66] ______, *Sums of two cubes*, Internat. Math. Res. Notices **4** (1995), 181–185.

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

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

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

[70] ______, *An explicit version of Birch's theorem*, Acta Arith. **85** (1998), 79–96.

[71] ______, *Sums of three cubes*, (submitted).

[72] ______, *On simultaneous additive equations IV*, Mathematika (to appear).

ABSTRACT

Exponential Sums and Diophantine Problems

by
Scott T. Parsell

Chair: Trevor D. Wooley

This work is concerned with the theory of exponential sums and their application to various Diophantine problems. Particular attention is given to exponential sums over smooth numbers, *i.e.* numbers having no large prime factors.

As an application of the theory of exponential sums in a single variable, we consider pairs of Diophantine inequalities of different degrees. Specifically, we show that two additive forms, one cubic and one quadratic, with real coefficients in at least 13 variables and satisfying suitable conditions, take arbitrarily small values simultaneously at integer points. In fact, we obtain a quantitative version of this result, which indicates how rapidly the forms can be made to approach zero as the size of the variables increases. Moreover, we obtain a lower bound for the density of integer points at which these small values occur.

We then proceed to study double exponential sums over smooth numbers by developing a version of the Vaughan-Wooley iterative method. We obtain estimates for mean values of these exponential sums, and these estimates are then used within the fabric of the Hardy-Littlewood method to obtain a lower bound for the density of rational lines on the hypersurface defined by an additive equation. We show that one obtains the expected density provided that the number of variables is sufficiently large in terms of the degree and that certain natural local solubility hypotheses are satisfied. We also consider applications to a two-dimensional generalization of Waring's problem and to fractional parts of polynomials in two variables.

Finally, we refine the above analysis in the case of a cubic hypersurface to show that the expected density of lines is obtained whenever the defining equation has at least 58 variables. In the process, we obtain a result on the paucity of non-trivial solutions to an associated system of Diophantine equations.