Acta Math., 162 (1989), 1–71

A new iterative method in Waring's problem

by

R. C. VAUGHAN

Imperial College
London, England, U.K.

Contents

1. Introduction

As is usual in Waring's problem we take $G(k)$ to be the smallest number s such that every sufficiently large natural number is the sum of at most s kth powers of natural numbers.

In this memoir we introduce a new iterative process to Waring's problem. We are thereby able to improve all previous upper bounds for $G(k)$ when $k\geqslant 5$.

Hitherto the best upper bounds for $G(k)$ for smaller $k\geqslant 4$ have been obtained by variants of the iterative method of Davenport (see [D3], [T3], [Va4] and [Va5]).

When $5\leqslant k\leqslant 8$ we obtain

THEOREM 1.1. *We have* $G(5)\leqslant 19$, $G(6)\leqslant 29$, $G(7)\leqslant 41$, $G(8)\leqslant 58$.

This may be compared with the respective bounds 21, 31, 45, and 62 contained in [Va4] and [Va5].

1–898282 *Acta Mathematica* 162. Imprimé le 8 mars 1989

The methods described here also improve the known upper bound in the case of biquadrates provided that an obviously necessary local condition is satisfied.

THEOREM 1.2. *Suppose that* $1 \leq r \leq 12$. *Then every sufficiently large natural number in the residue class r modulo* 16 *is the sum of at most* 12 *biquadrates.*

This compares with Theorem 2 of [Va4] in which 13 appears in place of the 12.

For cubes, although we are unable to reduce the known upper bound 7 for $G(3)$ (see [L], [W] and [Va6]) we are able to make progress with a quite closely related problem.

THEOREM 1.3. *Let* $\mathcal{N}(N)$ *denote the number of natural numbers not exceeding N which are the sum of three positive cubes. Then for* $N \geq 3$ *we have*

$$\mathcal{N}(N) \gg N^{\frac{11}{12}-\varepsilon}$$

where the implicit constant depends only on the positive number ε.

This can be compared with Theorem 6 of [Va6] in which a similar lower bound occurs but with the exponent 19/21 in place of 11/12, and the final theorem of Hooley [H] in which a conditional lower bound is given with 18/19 in place of 11/12.

It is perhaps not without some interest that our methods give a new lower bound in general for the number $\mathcal{N}_k(N)$ of different natural numbers not exceeding N which are the sum of three kth powers.

THEOREM 1.4. *Suppose that* $k \geq 4$ *and*

$$\alpha_k = \frac{3}{k} - \frac{1}{k^2}.$$

Then for $N \geq 3$ *we have*

$$\mathcal{N}_k(N) \gg N^{\alpha_k - \varepsilon}.$$

Previously the smallest known exponents α_k were

$$\alpha_4 = \frac{19}{28} \qquad \text{(Davenport [D1])},$$

$$\alpha_5 = \frac{5}{9}, \quad \alpha_6 = \frac{59}{126} \qquad \text{(Davenport [D2])},$$

$$\alpha_7 = \frac{65}{161}, \quad \alpha_8 = \frac{77}{216} \qquad \text{(Davenport's methods)},$$

$$\alpha_k = \frac{3}{k} - \frac{1}{k^2} - \frac{2}{k^3} \quad (k \geqslant 9) \qquad \text{(Davenport and Erdős [DE])}.$$

Consider the exponential sum

$$S(\alpha) = \sum_{x \leqslant P} e(\alpha x^k). \tag{1.1}$$

Then the methods based on Vinogradov's mean value theorem give

$$\sup_{\mathfrak{m}} |S(\alpha)| \ll P^{1-\sigma(k)+\varepsilon}. \tag{1.2}$$

where $\mathfrak{m}$ is the canonical set of minor arcs associated with the kth powers and where

$$4\sigma(k) k^2 \log k \to 1 \quad \text{as} \quad k \to \infty.$$

For an account of this see Chapter 5 of [Va2].

In many applications of (1.2) it is of no importance that the sum is over all the members of $[1, P] \cap \mathbf{Z}$. The next theorem shows that one can do considerably better than (1.2) when the sum is over a certain restricted but quite dense subset of $[1, P] \cap \mathbf{Z}$.

THEOREM 1.5. *Let $\mathcal{m}$ denote the set of real numbers α with the property that whenever $a \in \mathbf{Z}$, $q \in \mathbf{N}$, $(a, q) = 1$ and $|\alpha - a/q| \leqslant q^{-1} P^{1/2-k}$ one has $q > P^{1/2}$, and let*

$$\varrho(k) = \max_{\substack{s \in \mathbf{N} \\ s \geqslant 2}} \frac{1}{4s} \left(1 - (k-2)\left(1 - \frac{1}{k}\right)^{s-2}\right).$$

Then for each positive number ε there is a subset $\mathcal{B}$ of $[1, P] \cap \mathbf{N}$ such that

$$\operatorname{card} \mathcal{B} \gg P/\log P$$

and the exponential sum

$$T(\alpha) = \sum_{x \in \mathcal{B}} e(\alpha x^k)$$

satisfies

$$\sup_{\mathcal{m}} |T(\alpha)| \ll P^{1-\varrho(k)+\varepsilon}.$$

Moreover $4\varrho(k) k \log k \to 1$ as $k \to \infty$.

It is quite easy to construct examples in which for suitable subsets B of $[1,P]\cap\mathbf{Z}$ one has

$$\sup_{\mathfrak{m}}\left|\sum_{x\in\mathscr{B}} e(\alpha x^k)\right| \gg P^{1-1/k}. \tag{1.3}$$

See, for example, Lemma 4.4 of [Va2]. Thus the exponent $\varrho(k)$ is quite close to the best that one might hope to establish. For much of the purposes of this memoir the nature of $\mathscr{B}$ is not of great importance. Later we will see that a result of the same strength holds for a very natural set $\mathscr{B}$.

Vinogradov [Vi2] has shown that when k is large

$$G(k)<k(2\log k+4\log\log k+2\log\log\log k+13). \tag{1.4}$$

By a somewhat different method Karatsuba [K] has improved this by replacing the $4\log\log k+2\log\log\log k+13$ by $2\log\log k+12$ and increasing the domain of validity to $k>4000$. For smaller values of k, Balasubramanian and Mozzochi [BM], in an amalgam of an earlier version of Vinogradov's methods ([Vi1]) with those of Davenport and Erdős [DE], Vaughan [Va1] and Thanigasalam [T1], [T2], have shown that

$$G(k)\leqslant \frac{3\log k+\log 108}{\log\dfrac{k}{k-1}}-4. \tag{1.5}$$

The proof of (1.4) depends, in particular, on an application of Vinogradov's mean value theorem to give an estimate for a complicated exponential sum on the minor arcs. By using Theorem 1.5 instead we are able to establish

THEOREM 1.6. *Suppose that $k\geqslant 9$ and $\varrho(k)$ is as in Theorem* 1.5. *Then*

$$G(k)\leqslant 7+\min_{\substack{v\in\mathbf{N}\\ v\geqslant k-1}} 2\left(v+\left[\frac{k-2}{2\varrho(k)}\left(1-\frac{1}{k}\right)^{v}\right]\right).$$

Theorem 1.6 gives a smaller bound than any previously known when $k>13$, and as $k\to\infty$ it gives

$$G(k)<2k\left(\log k+\log\log k+1+\log 2+O\left(\frac{\log\log k}{\log k}\right)\right).$$

For intermediate values of k the method can be refined. Thus we are able to establish the following upper bounds for $G(k)$.

k	$F(k)$	k	$F(k)$	k	$F(k)$	k	$F(k)$
9	75	12	125	15	171	18	217
10	93	13	141	16	187	19	232
11	109	14	156	17	202	20	248

Table 1.1

THEOREM 1.7. *When* $9 \leq k \leq 20$ *we have* $G(k) \leq F(k)$ *where* $F(k)$ *is given by Table* 1.1.

This may be compared with the respective bounds 82 ([Va4]), 103 ([T3]), 119, 134, 150, 165, 181, 197, 213, 229, 245, 262 ([T2]).

We remark that a deeper treatment of the major arcs that arise in the proof of Theorem 1.7 enables one to replace the upper bound $F(k)$ for $G(k)$ by $F(k)-1$ when $k \neq 9$ or 15.

There is one aspect of the work in this memoir which may well have consequences outside additive number theory, namely the realisation of an estimate for exponential sums of the kind contained in Theorem 1.5. In many applications it may not be possible to accommodate the somewhat artificial set $\mathscr{B}$ that is provided by the proof of that theorem. However, by only slightly weakening the hypothesis it is possible to establish a conclusion of the same strength in which the set $\mathscr{B}$ is replaced by one of great familiarity in multiplicative number theory.

THEOREM 1.8. *Suppose that* $0<\delta<1/2k$, *let* $\mathfrak{m}$ *denote the set of real numbers* α *with the property that whenever* $a \in \mathbf{Z}$, $q \in \mathbf{N}$, $(a,q)=1$ *and*

$$|\alpha - a/q| \leq q^{-1}P^{1/2+\delta k-k}$$

one has $q>P^{1/2+\delta k}$, *and let*

$$\varrho(k) = \max_{\substack{s\in \mathbf{N}\\ s\geq 2}} \frac{1}{4s}\left(1-(k-2)\left(1-\frac{1}{k}\right)^{s-2}\right).$$

Further let $\mathscr{A}(P,R)$ *denote the set of natural numbers not exceeding* P *with no prime divisor exceeding* R. *Then for each positive number* ε *there is a positive number* η *such that whenever* $2 \leq R \leq P^{\eta}$ *the exponential sum*

$$S(\alpha) = \sum_{x\in \mathscr{A}(P,R)} e(\alpha x^k)$$

satisfies

$$\sup_{\mathfrak{m}}|S(\alpha)| \ll P^{1+\varepsilon}(P^{-\delta}+P^{-\varrho(k)}).$$

As in Theorem 1.5 we have

$$4\varrho(k)\,k\log k \to 1 \quad \text{as} \quad k\to\infty.$$

Moreover, for each fixed η we have

$$\operatorname{card}\mathscr{A}(P,P^{\eta}) \sim c_{\eta}P \quad \text{as} \quad P\to\infty$$

where c_{η} is a positive number.

The proof of Theorem 1.7 in § 9 shows *inter alia* that the exponent $\varrho(k)$ in Theorems 1.5 and 1.8 can be refined to one that is superior to that provided by Weyl's inequality whenever $k\geqslant 8$.

The methods of this memoir apply equally well to diagonal forms. Let $G^*(k)$ denote the least number t such that whenever $s\geqslant t$ the equation

$$c_1x_1^k+\ldots+c_sx_s^k = 0 \tag{1.6}$$

has a non-trivial solution in integers $x_1,\ldots,x_s$ when the coefficients $c_1,\ldots,c_s$

(i) are not all of the same sign when k is even and

(ii) are such that for every q (1.6) has a solution modulo q with $(x_j,q)=1$ for some j.

Then $G(k)$ may be replaced by $G^*(k)$ in each of the bounds. Thus it follows that $G^*(k)\leqslant k^2+1$ for all $k\geqslant 4$, in particular settling the stubborn case $k=10$, and so completing a programme initiated by Davenport and Lewis [DL]. The bound k^2+1 is of particular interest, since as Davenport and Lewis show, when $k+1$ is prime and $s=k^2$ there are $c_1,\ldots,c_s$ and a prime p for which (1.6) has no non-trivial p-adic solution.

After the seminal work of Hardy and Littlewood on additive number theory, and Waring's problem in particular (see [HL]), the best upper bounds for $G(k)$ when $k\geqslant 4$ have been based, in essence, on prior estimates for the number of solutions of auxiliary equations of the form

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

with the x_j and y_j lying in ranges of the kind

$$P_j < x_j < 2P_j, \quad P_j < y_j < 2P_j$$

where $P_1 \geqslant P_2 \geqslant \ldots$. The use of *diminishing ranges* in this context was refined and perfected by Davenport (see [D3]) and Vinogradov (see [Vi3] and [Va2]), and the recent work of Thanigasalam [T3] and Vaughan [Va4], [Va5] is largely based on a variant of the case $l=k-2$ of Theorem 1 of Davenport [D2].

The use of diminishing ranges in (1.7), whilst conferring a number of benefits, has one serious drawback, namely that the homogeneity of (1.7) is lost.

The underlying theme of this memoir is the conservation of homogeneity in equations such as (1.7). Thus we consider (1.7) with $x_j \in \mathscr{A}$, $y_j \in \mathscr{A}$ where $\mathscr{A}$ is a fairly dense subset of $[1, P] \cap \mathbf{Z}$. For a suitable $\mathscr{A}$ we relate the number of such solutions to the number of solutions of

$$x^k + m^k(z_2^k + \ldots + z_s^k) = y^k + m^k(t_2^k + \ldots + t_s^k) \tag{1.8}$$

with $x \leqslant P$, $y \leqslant P$, $M < m \leqslant M'$, $z_j \in \mathscr{B}$, $t_j \in \mathscr{B}$ where $\mathscr{B}$ has similar properties to $\mathscr{A}$, but $\mathscr{B} \subset [1, P/M] \cap \mathbf{Z}$. Then by the use of ideas stemming from the diminishing range circle of ideas combined with Hölder's inequality and the homogeneity of

$$z_2^k + \ldots + z_s^k - t_2^k - \ldots - t_s^k$$

we are able to estimate the number of solutions of (1.8) in terms of the number of solutions of (1.7) with s replaced by a number not exceeding s and with $\mathscr{A}$ replaced by $\mathscr{B}$. This enables an iterative procedure of an entirely new kind to be created. In a certain sense this does for a single equation what the arguments underlying the proof of Vinogradov's mean value theorem do for the corresponding system of equations.

It transpires that our technique puts no serious obstacle in the way of methods that have been developed in the context of diminishing ranges. Thus the technique has great flexibility.

An important role is played throughout this work by the set $\mathscr{A}(P, R)$ of natural numbers not exceeding P with no prime factor exceeding R. Other sets could be substituted in some of the arguments described herein, but no alternate seems to provide the same general degree of flexibility. This observation allied with a perusal of the methods used to establish Theorem 2 of [Va6] and Theorem 2 of [Va7] suggests that the greatest barrier to procuring improved estimates for the number of unrestricted solutions to (1.7) is the presence of x_j and y_j having large prime factors.

In § 2 we establish the basic relationships between the solutions of (1.7) and (1.8) for various choices of $\mathscr{A}$, $\mathscr{B}$ and s. In § 3 we analyse various exponential sums that arise. This leads to another relationship between (1.7) and (1.8) that is particularly effective when s is large.

The conclusions of §§ 2 and 3 are applied in § 4 to obtain general bounds for the number of solutions of (1.7) and then Theorems 1.1 and 1.2 are established in § 5 and Theorems 1.3 and 1.4 in § 6.

A simplified bound for the number of solutions of (1.7) is given in Theorem 7.1. This is used to establish Theorem 1.5 in § 7 and Theorem 1.6 in § 8.

In § 9 various methods from earlier sections are brought together to establish Theorem 1.7, and finally Theorem 1.8 established in § 10.

1.1. Notation

In general we use upper case Latin letters to denote real numbers which exceed 2 unless otherwise stated, lower case Latin letters to denote integers and lower case Greek letters to denote positive real numbers. In particular $k, m, n, q, r, s, t \in \mathbb{N}$ with $k \geqslant 3$, and p denotes a prime number. Implicit constants may depend on k, s, t, ε. Throughout we think of k as being fixed. Therefore in explicit constants such as the D_r and $C_r(\varepsilon)$ of Theorem 4.1, and the $P_0(\eta, s)$ of Theorem 4.2, the k is suppressed.

2. The reduction of the auxiliary equation

Let

$$\mathscr{A}(P,R) = \{n : n \leqslant P, p|n \Rightarrow p \leqslant R\}, \tag{2.1}$$

let $S_s(P,R)$ denote the number of solutions of

$$x_1^k + \ldots + x_s^k = y_1^k + \ldots + y_s^k \tag{2.2}$$

with

$$x_j \in \mathscr{A}(P,R), \quad y_j \in \mathscr{A}(P,R), \tag{2.3}$$

and for a given real number θ with $0<\theta<1$ let $T_s(P,R,\theta)$ denote the number of solutions of

$$x^k + m^k(x_1^k + \ldots + x_{s-1}^k) = y^k + m^k(y_1^k + \ldots + y_{s-1}^k) \tag{2.4}$$

with

$$x \leqslant P, \quad y \leqslant P, \quad x \equiv y \pmod{m^k}, \quad P^\theta < m \leqslant \min(P, P^\theta R), \tag{2.5}$$

$$x_j \in \mathscr{A}(P^{1-\theta}, R), \quad y_j \in \mathscr{A}(P^{1-\theta}, R). \tag{2.6}$$

The lemma below relates S_s to T_s.

LEMMA 2.1. *Let $\theta=\theta(s,k)$ satisfy $0<\theta<1$ and suppose that*

$$1 \leqslant D \leqslant P.$$

Then

$$S_s(P,R) \ll \Big(\sum_{d>D} (S_s(P/d,R))^{1/s}\Big)^s + S_s(D^{1-\theta}P^\theta, R) + P^\varepsilon\Big(\sum_{d\leqslant D}((P/d)^\theta R)^{2-3/s}(T_s(P/d,R,\theta))^{1/s}\Big)^s.$$

When $s>k$ and R is not too small by comparison with P we expect that $S_s(P,R) \gg P^\sigma$ and $T_s(P,R,\theta) \gg P^\tau$ with $\sigma>s$, $\tau>s$. Thus for a suitable choice of D the first two terms on the right of the above inequality can be expected to be small compared with the left hand side and the third term will be dominated by the term in the sum with $d=1$. Thus in principle the lemma says that either

$$S_s(P,R) \ll P^s$$

or

$$S_s(P,R) \ll (P^\theta R)^{2s-3} T_s(P,R,\theta).$$

Proof of Lemma 2.1. For a given solution of (2.2) satisfying (2.3) let

$$d_j = (x_j, y_j) \quad (1 \leqslant j \leqslant s).$$

Now let S' denote the number of those solutions for which $d_j>D$ for at least one j, let S'' denote the number for which

$$d_j \leqslant D \tag{2.7}$$

for every j and

$$\max\{x_j, y_j\} \leqslant d_j^{1-\theta} P^\theta \tag{2.8}$$

for at least one j, and let S''' denote the number for which $d_j \leq D$ for every j and (2.8) holds for *no* j. Then

$$S_s(P,R) \leq 3\max\{S', S'', S'''\}.$$

First suppose that $S' \geq \max\{S'', S'''\}$, so that

$$S_s(P,R) \leq 3S'.$$

Let

$$f(\alpha; Q, R) = \sum_{x \in \mathcal{A}(Q,R)} e(\alpha x^k). \tag{2.9}$$

Then

$$S' \ll \sum_{d>D} \int_0^1 |f(\alpha d^k; P/d, R)^2 f(\alpha; P, R)^{2s-2}| \, d\alpha.$$

Hence, by Hölder's inequality,

$$S_s(P,R) \ll \sum_{d>D} (S_s(P/d,R))^{1/s} (S_s(P,R))^{1-1/s}$$

and the lemma now follows in the first case.

Secondly suppose that $S'' \geq \max\{S''', S'\}$, so that

$$S_s(P,R) \leq 3S''.$$

Then for a solution counted by S'' we have (2.7) for all j and (2.8) for some j, say $j=i$. Thus

$$d_i \leq D \quad \text{and} \quad \max\{x_i, y_i\} \leq d_i^{1-\theta} P^\theta,$$

so that

$$\max\{x_i, y_i\} \leq D^{1-\theta} P^\theta.$$

Hence

$$S'' \ll \int_0^1 |f(\alpha; D^{1-\theta}P^\theta, R)^2 f(\alpha; P, R)^{2s-2}| \, d\alpha.$$

Therefore, by Hölder's inequality once more, we have

$$S_s(P,R) \ll (S_s(D^{1-\theta}P^{\theta},R))^{1/s}(S_s(P,R))^{1-1/s}$$

and so the result ensues in the second case.

Lastly suppose that $S''' \geqslant \max\{S', S''\}$, so that

$$S_s(P,R) \leqslant 3S'''. \tag{2.10}$$

Then for a given solution of (2.2) counted by S''' we have, for every j,

$$d_j \in \mathscr{A}(D,R) \quad \text{and} \quad \max\{x_j, y_j\} > d_j^{1-\theta}P^{\theta}.$$

Let $u_j = x_j/d_j$, $v_j = y_j/d_j$, so that

$$(u_j, v_j) = 1 \quad \text{and} \quad \max\{u_j, v_j\} > (P/d_j)^{\theta},$$

and let m_j denote the smallest divisor of $\max\{u_j, v_j\}$ exceeding $(P/d_j)^{\theta}$. Since none of the prime divisors of $\max\{u_j, v_j\}$ exceed R we have

$$(P/d_j)^{\theta} < m_j \leqslant \min(P/d_j, (P/d_j)^{\theta}R). \tag{2.11}$$

Thus

$$S''' \ll \sum_{\eta_1} \dots \sum_{\eta_s} S^{\mathrm{IV}}(\eta_2, \dots, \eta_s) \tag{2.12}$$

where the summation is over $\eta_1, \dots, \eta_s$ with $\eta_j = \pm 1$ and where $S^{\mathrm{IV}}(\eta_1, \dots, \eta_s)$ is the number of solutions of

$$\sum_{j=1}^{s} \eta_j d_j^k (x_j^k - m_j^k y_j^k) = 0$$

with

$$d_j \in \mathscr{A}(D,R),$$

$$x_j \in \mathscr{A}(P/d_j, R), \quad (x_j, m_j) = 1, \quad y_j \in \mathscr{A}(P/(d_j m_j), R)$$

and m_j satisfying (2.11).

Let

$$f_m(\alpha; Q, R) = \sum_{\substack{x \in \mathscr{A}(Q,R) \\ (x,m)=1}} e(\alpha x^k). \tag{2.13}$$

$$F_j(\alpha) = f_{m_j}(\eta_j d_j^k \alpha; P/d_j, R) f(-\eta_j d_j^k m_j^k \alpha; P/(d_j m_j), R).$$

Then

$$S^{\mathrm{IV}}(\eta_1, \ldots, \eta_s) \leqslant \int_0^1 \prod_{j=1}^s \left(\sum_{d_j \in \mathscr{A}(D,R)} \sum_{m_j}{}' F_j(\alpha) \right) d\alpha$$

where Σ'_{m_j} denotes summation over m_j, satisfying (2.11).

Let

$$X_j(\alpha) = |f_{m_j}(d_j^k \alpha; P/d_j, R)^2 f(d_j^k m_j^k \alpha; P/(d_j m_j), R)^{2s-2}| \tag{2.14}$$

and

$$Y(\alpha) = \left| \prod_{j=1}^s f_{m_j}(d_j^k \alpha; P/d_j, R) \right|. \tag{2.15}$$

Then, by (2.12),

$$S''' \ll \sum_{\substack{d_1 \\ d_j \in \mathscr{A}(D,R)}} \cdots \sum_{d_s} \sum_{m_1}{}' \cdots \sum_{m_s}{}' \int_0^1 Y(\alpha)^{\frac{s-2}{s-1}} \prod_{j=1}^s \left(X_j(\alpha)^{\frac{1}{2s-2}} \right) d\alpha. \tag{2.16}$$

By Hölder's inequality we have

$$\int_0^1 Y(\alpha)^{\frac{s-2}{s-1}} \prod_{j=1}^s \left(X_j(\alpha)^{\frac{1}{2s-2}} \right) d\alpha \leqslant \left(\int_0^1 Y(\alpha)^2 d\alpha \right)^{\frac{s-2}{2s-2}} \prod_{j=1}^s \left(\int_0^1 X_j(\alpha)\, d\alpha \right)^{\frac{1}{2s-2}}$$

and by (2.9), (2.13) and (2.15), and by considering the underlying diophantine equation, we have

$$\int_0^1 Y(\alpha)^2\, d\alpha \leqslant \int_0^1 Z(\alpha)^2\, d\alpha$$

where

$$Z(\alpha) = \left| \prod_{j=1}^{s} f(d_j^k \alpha; P/d_j, R) \right|. \tag{2.17}$$

Therefore, by Hölder's inequality and (2.14),

$$\begin{aligned} &\sum_{m_1}{}' \cdots \sum_{m_s}{}' \int_0^1 Y(\alpha)^{\frac{s-2}{s-1}} \prod_{j=1}^{s} \left(X_j(\alpha)^{\frac{1}{2s-2}} \right) d\alpha \\ &\ll \left(\int_0^1 Y(\alpha)^2 \, d\alpha \right)^{\frac{s-2}{2s-2}} \left(\sum_{m_1}{}' \cdots \sum_{m_s}{}' 1 \right)^{\frac{2s-3}{2s-2}} \left(\sum_{m_1}{}' \cdots \sum_{m_s}{}' \prod_{j=1}^{s} \int_0^1 X_j(\alpha) \, d\alpha \right)^{\frac{1}{2s-2}} \\ &\ll \left(\int_0^1 Z(\alpha)^2 \, d\alpha \right)^{\frac{s-2}{2s-2}} \left(\prod_{j=1}^{s} ((P/d_j)^\theta R)^{2s-3} U(P/d_j, R, \theta) \right)^{\frac{1}{2s-2}} \end{aligned}$$

where

$$\left. \begin{array}{c} U(Q, R, \theta) \text{ is the number of solutions of (2.4) with} \\ x \leqslant Q, \quad y \leqslant Q, \quad (xy, m) = 1, \quad Q^\theta < m \leqslant \min(Q, Q^\theta R), \\ x_j \in \mathscr{A}(Q^{1-\theta}, R), \quad y_j \in \mathscr{A}(Q^{1-\theta}, R). \end{array} \right\} \tag{2.18}$$

Therefore, by (2.16) and Hölder's inequality,

$$S''' \ll \left(\sum_{\substack{d_1 \\ d_j \in \mathscr{A}(D,R)}} \cdots \sum_{d_s} \int_0^1 Z(\alpha)^2 \, d\alpha \right)^{\frac{s-2}{2s-2}} \left(\sum_{\substack{d_1 \\ d_j \in \mathscr{A}(D,R)}} \cdots \sum_{d_s} \prod_{j=1}^{s} V(d_j)^{1/s} \right)^{\frac{s}{2s-2}}$$

where

$$V(d) = ((P/d)^\theta R)^{2s-3} U(P/d, R, \theta). \tag{2.19}$$

By (2.17)

$$\sum_{\substack{d_1 \\ d_j \in \mathscr{A}(D,R)}} \cdots \sum_{d_s} \int_0^1 Z(\alpha)^2 \, d\alpha$$

is the number of solutions

$$d_1^k x_1^k + \ldots + d_s^k x_s^k = d_1^k y_1^k + \ldots + d_s^k y_s^k$$

with $d_j \in \mathcal{A}(D,R)$, $x_j \in \mathcal{A}(P/d_j,R)$, $y_j \in \mathcal{A}(P/d_j,R)$. Hence it it

$$\ll P^{\varepsilon} S_s(P,R).$$

Therefore, by (2.10),

$$S_s(R,P) \ll P^{\varepsilon}\Bigg(\sum_{d\in\mathcal{A}(D,R)} V(d)^{1/s}\Bigg)^s.$$

Hence, in view of (2.19) it remains to show that

$$U(Q,R,\theta) \ll Q^{\varepsilon} T_s(Q,R,\theta). \tag{2.20}$$

For a given m let $\mathcal{B}(u)$ denote the set of solutions of the congruence

$$z^k \equiv u \pmod{m^k}.$$

Then

$$\operatorname{card} \mathcal{B}(u) \ll m^{\varepsilon} \quad ((u,m)=1). \tag{2.21}$$

Clearly in (2.4), $x^k \equiv y^k \pmod{m^k}$. Thus each solution of (2.4) can be classified according to the common residue class modulo m^k of x^k and y^k. Let

$$g_m(\alpha,z) = \sum_{\substack{x\le Q \\ x\equiv z \pmod{m^k}}} e(\alpha x^k).$$

Then, by (2.18),

$$U(Q,R,\theta) \le \sum_{Q^{\theta}<m\le \min(Q,Q^{\theta}R)} U_m$$

where

$$U_m = \int_0^1 G_m(\alpha)\,|f(m^k\alpha;Q,R)|^{2s-2}\,d\alpha$$

and

$$G_m(\alpha) = \sum_{\substack{u=1 \\ (u,m)=1}}^{m^k} \Bigg|\sum_{z\in\mathcal{B}(u)} g_m(\alpha,z)\Bigg|^2.$$

Hence, by Cauchy's inequality and (2.21),

$$G_m(\alpha) \ll m^\varepsilon \sum_{\substack{u=1 \\ (u,m)=1}}^{m^k} \sum_{z \in \mathscr{B}(u)} |g_m(\alpha, z)|^2 = m^\varepsilon \sum_{\substack{z=1 \\ (z,m)=1}}^{m^k} |g_m(\alpha, z)|^2.$$

Therefore we have (2.20) as required.

This completes the proof of Lemma 2.1.

Henceforward we shall suppose that θ satisfies

$$0 < \theta \leqslant \frac{1}{k} \tag{2.22}$$

and put

$$M = P^\theta, \quad H = P^{1-k\theta}, \quad Q = P^{1-\theta}. \tag{2.23}$$

We shall normally suppose that P is large and that R is at most a fairly small power of P, so that in particular

$$R^{2k-2} \leqslant P^{k-5}M \ (k \geqslant 5) \quad \text{and} \quad 2MR \leqslant P^{1/2}. \tag{2.24}$$

Consider equation (2.4). We put $z=x+y$ and $h=(x-y)m^{-k}$. Thus $2x=z+hm^k$ and $2y=z-hm^k$. Hence, by (2.5),

$$T_s(P, R, \theta) \leqslant U_0 + 2U_1, \tag{2.25}$$

where U_0 is the number of solutions of (2.4) with (2.5), (2.6) and $x=y$, and U_1 is the number of solutions of

$$(z+hm^k)^k + (2m)^k(x_i^k + \ldots + x_{s-1}^k) = (z-hm^k)^k + (2m)^k(y_1^k + \ldots + y_{s-1}^k) \tag{2.26}$$

with

$$z \leqslant 2P, \quad h \leqslant H, \quad M < m \leqslant MR, \quad x_j \in \mathscr{A}(Q, R), \quad y_j \in \mathscr{A}(Q, R). \tag{2.27}$$

Obviously

$$U_0 \leqslant PMRS_{s-1}(Q, R). \tag{2.28}$$

We now wish to relate U_1 to $S_s(Q, R)$ and $S_{s-1}(Q, R)$. One line of attack is through an argument of Davenport [D2]. However the homogeneity of the x_i and y_i in (2.26) enables significant improvements to be made.

Let Δ_1 denote the forward difference operator

$$\Delta_1(f(x);h)=f(x+h)-f(x) \tag{2.29}$$

and define Δ_j recursively by

$$\Delta_{j+1}(f(x);h_1,\ldots,h_{j+1})=\Delta_1(\Delta_j(f(x);h_1,\ldots,h_j);h_{j+1}). \tag{2.30}$$

Now let

$$\Psi_j=\Psi_j(z;h,h_2,\ldots,h_j,m)=m^{-k}\Delta_j(f(z);2hm^k,h_2,\ldots,h_j) \tag{2.31}$$

where $f(z)=(z-hm^k)^k$, let $R(n)$ denote the number of solutions of

$$\Psi_j=n \tag{2.32}$$

with

$$m\leqslant MR,\quad h\leqslant H,\quad h_i\leqslant 2P,\quad z\leqslant 2P \tag{2.33}$$

and let

$$N_j=\sum_n R(n)^2. \tag{2.34}$$

We now introduce the exponential sum

$$F_j(\alpha)=\sum_{M<m\leqslant MR}\sum_{h\leqslant H}\sum_{h_2\leqslant 2P}\cdots\sum_{h_j\leqslant 2P}\sum_{z\in\mathscr{B}}e(\alpha\Psi_j) \tag{2.35}$$

where $\mathscr{B}=\mathscr{B}(h_2,\ldots,h_j)$ is the set of z satisfaying $0<z\leqslant 2P-h_2-\ldots-h_j$. Thus, by (2.9), (2.26) and (2.31),

$$U_1\leqslant\int_0^1 F_1(\alpha)\,|f(2^k\alpha;Q,R)|^{2s-2}\,d\alpha.$$

Hence, by (2.25) and (2.28),

$$T_s(P,R,\theta)\ll PMRS_{s-1}(Q,R)+\int_0^1 F_1(\alpha)\,|f(2^k\alpha;Q,R)|^{2s-2}\,d\alpha. \tag{2.36}$$

By the standard Weyl technique for estimating exponential sums (see Lemma 2.3 of [Va2]) we have

$$F_1(\alpha) \ll P^{1-2J}HMR + P^{1-2jJ}(HMR)^{1-2J}|F_j(\alpha)|^{2J} \tag{2.37}$$

where the implicit constant depends at most on j and where

$$J = 2^{-j}. \tag{2.38}$$

The next lemma relates T_s to S_s and S_{s-1} and is particularly useful when $S_{s-1} \approx P^\lambda$ with λ large compared with $2s-2-k$.

LEMMA 2.2. *Suppose that $\theta=\theta(s,k)$ satisfies (2.22), and j satisfies $1 \leqslant j \leqslant k-1$ and $2^j \geqslant s$. Then*

$$T_s(P,R,\theta) \ll (PMR + P^{1-2J}HMR)\, S_{s-1}(Q,R) + (HMR)^{1-2J}P^{1-2jJ}N_j^J\, S_{s-1}(Q,R)^{1-sJ} S_s(Q,R)^{(s-1)J}.$$

Proof. By (2.34) and (2.35),

$$\int_0^1 |F_j(\alpha)|^2\, d\alpha \leqslant N_j.$$

The lemma now ensues from (2.36) and (2.37) *via* Hölder's inequality.

We require an estimate for N_j. By (2.29) and (2.30) we have

$$\Delta_j(x^k, h_1, \ldots, h_j) = \sum_{\theta_1=\pm 1} \ldots \sum_{\theta_j=\pm 1} \theta_1 \ldots \theta_j \left(x + \frac{1+\theta_1}{2} h_1 + \ldots + \frac{1+\theta_j}{2} h_j \right)^k.$$

By writing $\xi = 2x + h_1 + \ldots + h_j$ and employing the multinomial theorem this becomes

$$\sum_{\theta_1=\pm1} \ldots \sum_{\theta_j=\pm1} \sum_{u_0 \geqslant 0} \ldots \sum_{\substack{u_j \geqslant 0 \\ u_0+\ldots+u_j=k}} \frac{k!\,\theta_1 \ldots \theta_j \xi^{u_0}}{u_0! \ldots u_j!\, 2^k} (\theta_1 h_1)^{u_1} \ldots (\theta_j h_j)^{u_j}$$

$$= \sum_{u \geqslant 0} \sum_{v_1 \geqslant 0} \ldots \sum_{\substack{v_j \geqslant 0 \\ u+2v_1+\ldots+2v_j=k-j}} \frac{k!\, 2^{j-k} h_1 \ldots h_j \xi^u}{u!(2v_1+1)! \ldots (2v_j+1)!} h_1^{2v_1} \ldots h_j^{2v_j}.$$

Therefore, by (2.31),

$$\Psi_j = \sum_{u \geqslant 0} \sum_{v_1 \geqslant 0} \ldots \sum_{\substack{v_j \geqslant 0 \\ u+2v_1+\ldots+2v_j=k-j}} \frac{k!\, 2^{1+j-k} h h_2 \ldots h_j}{u!(2v_1+1)! \ldots (2v_j+1)!} \xi^u (2hm^k)^{2v_1} h_2^{2v_2} \ldots h_j^{2v_j} \tag{2.39}$$

where

$$\xi = 2z + h_2 + \ldots + h_j. \tag{2.40}$$

In particular

$$\Psi_{k-2} = \frac{k!}{12} hh_2 \ldots h_{k-2}(3\xi^2 + (2hm^k)^2 + h_2^2 + \ldots + h_{k-2}^2). \tag{2.41}$$

LEMMA 2.3. *Suppose that $j \leqslant k-1$. Then*

$$N_j \ll P^{j+\varepsilon} HM^2R^2.$$

Moreover, if $k-j$ is odd and $j \leqslant k-3$, or $j=k-2$, or $j=k-4$, then

$$N_j \ll P^{j+\varepsilon} HMR.$$

Proof. By (2.39)

$$\Psi_j = 2^{1+j-k} hh_2 \ldots h_j \Omega(\xi; 2hm^k, h_2, \ldots, h_j)$$

where Ω is a polynomial of degree $k-j$ in ξ with integer coefficients and leading coefficient $k!/(k-j)!$. Thus, for each given m the equation (2.32) has $\ll n^\varepsilon$ solutions in $h, h_2, \ldots, h_j, \xi$, and hence, by (2.40), in $h, h_2, \ldots, h_j, z$. The first part of the lemma now follows from (2.34).

When $k-j$ is odd, and $j \leqslant k-3$, (2.39) gives

$$\Psi_j = 2^{1+j-k} hh_2 \ldots h_j \xi \Omega(2hm^k; \xi, h_2, \ldots, h_j)$$

where Ω is a polynomial of degree $k-j-1$ in $2hm^k$ with integer coefficients and leading coefficient $k!/(k-j)!$. Hence the equation (2.32) has $\ll n^\varepsilon$ solutions in $h, h_2, \ldots, h_j, \xi, m$ and so in $z, m, h, h_2, \ldots, h_j$.

The number of solutions of

$$3x^2 + y^2 = m$$

in integers x, y is $O(m^\varepsilon)$ (see, for example [E]). Hence, by (2.41), when $j=k-2$ the number of solutions of (2.32) in $z, m, h, h_2, \ldots, h_j$ is again $\ll n^\varepsilon$.

Now suppose that $j=k-4$. Then, by (2.39),

$$\Psi_j = \frac{k!}{8640} hh_2 \ldots h_{k-4}(\Theta^2 + \Theta\Xi - \Xi^2 + \Phi)$$

where

$$\Theta = 6\xi^2+6(2hm^k)^2+6h_2^2+\dots+6h_{k-4}^2, \quad \Xi = 3\xi^2-3(2hm^k)^2-2h_2^2-\dots-2h_{k-4}^2$$

and Φ is a form in $h_2, \dots, h_{k-4}$ of degree 4 with integer coefficients. By (2.39) and (2.4), $\Psi_j>0$ for all choices of $z, h, m, h_2, \dots, h_{k-4}$ satisfying (2.33). Moreover, by the theory of $\mathbf{Q}(\sqrt{5})$ (or see [E]), the number of solutions of

$$x^2+xy-y^2 = b$$

in integers x, y is $\ll 1+|b|^\varepsilon$. Hence the number of solutions of (2.32) in $z, h, m, h_2, \dots, h_{k-4}$ is $\ll n^\varepsilon$ in this case also.

This concludes the proof of the lemma.

3. A variation on the main theme

By (2.35) and (2.41), and Cauchy's inequality, we have

$$|F_{k-2}(\alpha)|^2 \leq D(\alpha)\, E(\alpha) \tag{3.1}$$

where

$$D(\alpha) = \sum_{h\leq H}\sum_{h_2\leq 2P}\dots\sum_{h_{k-2}\leq 2P}\left|\sum_{z\in\mathcal{B}} e\left(\frac{1}{4}\alpha k!\, hh_2\dots h_{k-2}\xi^2\right)\right|^2 \tag{3.2}$$

and

$$E(\alpha) = \sum_{h\leq H}\sum_{h_2\leq 2P}\dots\sum_{h_{k-2}\leq 2P}\left|\sum_{M<m\leq MR} e\left(\frac{1}{3}\alpha k!\, h^3h_2\dots h_{k-2}m^{2k}\right)\right|^2. \tag{3.3}$$

As an alternative to the estimation given in the previous section of the integral of the right of (2.36) we use a form of the Hardy-Littlewood method. Thus we require estimates for F_{k-2} that depend on the nature of the rational approximations to α. This is most readily accomplished through estimates for D and E. The first of these two exponential sums can be estimated quite easily.

Note that throughout this section implicit constants depend at most on k and ε.

LEMMA 3.1. *Suppose that $(a, q)=1$ and $|\alpha-a/q|\leq q^{-2}$. Then*

$$D(\alpha) \ll P^\varepsilon\left(\frac{P^{k-1}H}{q+Q^k|\alpha q-a|}+P^{k-2}H+q+Q^k|\alpha q-a|\right).$$

Proof. We square out the innermost sum in (3.2). This gives a double sum over z_1 and z_2, say. We put $h_{k-1}=z_1-z_2$ and $j=k!\,2hh_2\dots h_{k-2}|h_{k-1}|$, and sum over z_1. By (2.40) we obtain

$$D(\alpha)\ll HP^{k-2}+P^{\varepsilon}\sum_{1\leqslant j\leqslant 2k!\,H(2P)^{k-2}}\min(P,\|\alpha_j\|^{-1}),$$

where $\|\theta\|$ denotes the distance of θ from a nearest integer. Thus, by Lemma 2.2 of [Va2], if $|\alpha-b/r|\leqslant r^{-2}$ with $(b,r)=1$, then

$$D(\alpha)\ll P^{\varepsilon}(HP^{k-1}r^{-1}+HP^{k-2}+r). \tag{3.4}$$

When $Q^k|\alpha q-a|\leqslant q$ the lemma is immediate on taking $b=a$, $r=q$.

When $Q^k|\alpha q-a|>q$, choose b, r so that $(b,r)=1$, $r\leqslant 2/|\alpha q-a|$, $|\alpha-b/r|\leqslant|\alpha q-a|/(2r)$. If $b/r=a/q$, then $0<|\alpha-a/q|\leqslant|\alpha-a/q|/2$ which is impossible. Hence

$$\left|\frac{b}{r}-\frac{a}{q}\right|\geqslant\frac{1}{qr}.$$

Therefore $|\alpha-a/q|\geqslant 1/(qr)-|\alpha-b/r|\geqslant 1/(qr)-|\alpha q-a|/(2r)\geqslant 1/(2qr)$. Thus

$$\frac{1}{r}\leqslant 2|\alpha q-a|.$$

Therefore, by (3.4),

$$D(\alpha)\ll P^{\varepsilon}(HP^{k-1}|\alpha q-a|+HP^{k-2}+|\alpha q-a|^{-1}).$$

Moreover, by (2.23), $HP^{k-1}=Q^k$. This gives the lemma.

The estimation of E is harder, and requires special arguments when $k=3$ or 4. We first of all treat the case $k\geqslant 5$.

LEMMA 3.2. *Suppose that $k\geqslant 5$, that* (2.24) *holds, that $M^k\leqslant X\leqslant Q^kM^{-k}$, and that $(a,q)=1$, $q\leqslant X$ and $|\alpha-a/q|\leqslant q^{-1}X^{-1}$. Then*

$$E(\alpha)\ll\frac{P^{k-3+\varepsilon}HM^2R^2}{(q+Q^k|\alpha q-a|)^{1/k}}+P^{k-3+\varepsilon}HMR^2.$$

Proof. By (3.3), we obtain

$$E(\alpha)\ll\sum_{h\leqslant H}\sum_{j\leqslant k!(2P)^{k-3}}P^{\varepsilon}\left|\sum_{M<m\leqslant MR}e(\alpha_jh^3m^{2k})\right|^2\ll P^{k-3+\varepsilon}HMR+P^{\varepsilon}E_0,$$

on squaring out and summing over j, where

$$E_0 = \sum_{h \leqslant H} \sum_{m_1} \sum_{m_2}_{M < m_1 < m_2 \leqslant MR} \min(P^{k-3}, \|\alpha h^3(m_2^{2k} - m_1^{2k})\|^{-1}).$$

Choose b, r with $r \leqslant 2H^3(MR)^{2k}$, $(b, r) = 1$, $|\alpha - b/r| \leqslant 1/(2H^3(MR)^{2k} r)$. Then for $h \leqslant H$ and $M < m_1 < m_2 \leqslant MR$ we have

$$|(\alpha - b/r)\, h^3(m_2^{2k} - m_1^{2k})| \leqslant 1/(2r).$$

Thus

$$\|\alpha h^3(m_2^{2k} - m_1^{2k})\| \geqslant \|bh^3(m_2^{2k} - m_1^{2k})/r\| - 1/(2r) \geqslant \frac{1}{2} \|bh^3(m_2^{2k} - m_1^{2k})/r\|$$

unless $r | (m_2^{2k} - m_1^{2k})$, in which case

$$\|\alpha h^3(m_2^{2k} - m_1^{2k})\| = h^3(m_2^{2k} - m_1^{2k})\, |\alpha - b/r|.$$

Hence

$$E_0 \ll E_1 + E_2$$

where

$$E_1 = \sum_{h \leqslant H} \sum_{\substack{M < m_1 < m_2 \leqslant MR \\ r \nmid h^3(m_2^{2k} - m_1^{2k})}} \sum \|bh^3(m_2^{2k} - m_1^{2k})/r\|^{-1}$$

and

$$E_2 = \sum_{h \leqslant H} \sum_{\substack{M < m_1 < m_2 \leqslant MR \\ r | h^3(m_2^{2k} - m_1^{2k})}} \sum \min\left(P^{k-3}, \left(h^3 M^{2k-1}(m_2 - m_1) \left|\alpha - \frac{b}{r}\right|\right)^{-1}\right).$$

When $h^3(m_2^{2k} - m_1^{2k}) = j$ each of h, $m_2^k - m_1^k$, $m_2^k + m_1^k$ is a divisor of j, and thus the number of solutions in h, m_2, m_1 is $\ll j^{\varepsilon}$. Therefore

$$E_1 \ll P^{\varepsilon} \sum_{\substack{j \leqslant H^3(MR)^{2k} \\ r \nmid j}} \|bj/r\|^{-1} \ll P^{\varepsilon}(H^3(MR)^{2k} r^{-1} + 1) r \log 2r$$

and so

$$E_1 \ll P^{\varepsilon} H^3 (MR)^{2k}.$$

Now

$$H^3(MR)^{2k} = P^{k-3} HMR^2 P^{5-k} M^{-1} R^{2k-2} \leqslant P^{k-3} HMR^2$$

by (2.24). Therefore we can concentrate on E_2.

In E_2 we write $(m_2, m_1)=j$, $n_i=m_i/j$, so that $j \leqslant MR$, $(n_2, n_1)=1$, $M/j<n_1<n_2 \leqslant MR/j$, $r \mid h^3 j^{2k}(n_2^{2k}-n_1^{2k})$ and $m_2-m_1=j(n_2-n_1)$.

For a given $h \leqslant H$ we put $(r, h^3)=e$ and write $e=e_1e_2^2e_3^3$ where e_3^3 is the largest cube divisor of e and e_2^2 is the largest square divisor of e/e_3^3. Hence $e_1e_2e_3 \mid h$. Let $h_0 = h/(e_1e_2e_3)$, so that

$$\left(\frac{r}{e}, h_0^3 e_1^2 e_2\right) = 1 \quad \text{and} \quad \frac{r}{e} \,\Big|\, j^{2k}(n_2^{2k}-n_1^{2k}).$$

Now given $h \leqslant H$ and $j \leqslant MR$ we put similarly $(r/e, j^{2k}) = f = f_1 f_2^2 \dots f_{2k}^{2k}$, so that $j = j_0 f_1 \dots f_{2k}$ for a suitable j_0, and

$$\left(\frac{r}{ef}, j_0^{2k} f_1^{2k-1} \dots f_{2k-1}\right) = 1 \quad \text{and} \quad \frac{r}{ef} \,\Big|\, (n_2^{2k}-n_1^{2k}).$$

Let $g=r/(ef)$ and put $n=n_1$, $l=n_2-n_1$, $g_0=(l, g)$, $l_0=l/g_0$, $g_1=g/g_0$. Thus

$$E_2 \ll \sum_{\substack{e \\ efg=r}} \sum_{f} \sum_{g} E_3$$

with

$$E_3 = \sum_{h_0 \leqslant H/(e_1e_2e_3)} \; \sum_{j_0 \leqslant MR/(f_1 \dots f_{2k})} E_4$$

and

$$E_4 = \sum_{\substack{g_0 \\ g_0 g_1 = g}} \sum_{g_1} \sum_{l_0} \sum_{n} \min\left(P^{k-3}, \left(h_0^3 e_1^3 e_2^3 e_3^3 M^{2k-1} j_0 f_1 \dots f_{2k} l_0 g_0 \left|\alpha - \frac{b}{r}\right|\right)^{-1}\right)$$

and with l_0 and n satisfying

$$(l_0, g_1) = 1, \quad l_0 \leqslant MR/(g_0 j_0 f_1 \dots f_{2k}),$$
$$M/(j_0 f_1 \dots f_{2k}) < n \leqslant MR/(j_0 f_1 \dots f_{2k}),$$
$$(n, n+l_0 g_0) = 1,$$
$$(n+l_0 g_0)^{2k} \equiv n^{2k} \pmod g.$$

The last two conditions imply that $(n, g)=1$. For a given n satisfying the last three conditions, choose x so that $nx\equiv 1 \pmod{g_1}$. Obviously this establishes a bijection between the residue class of n modulo g_1 and that of x. We also have

$$((n+l_0 g_0)^{2k}-n^{2k})\, g_0^{-1} \equiv 0 \pmod{g_1}.$$

Hence

$$((1+l_0 g_0 x)^{2k}-1)\, g_0^{-1} \equiv 0 \pmod{g_1},$$

whence

$$(1+l_0 g_0 x)^{2k} \equiv 1 \pmod g.$$

Now the congruence $y^{2k}\equiv 1 \pmod g$ has v solutions modulo g, say $y_1, \dots, y_v$, where $v \ll g^{\varepsilon}$. Hence $1+l_0 g_0 x \equiv y_i \pmod g$ for some $i \leqslant v$. Thus $y_i \equiv 1 \pmod{g_0}$ and $l_0 x \equiv (y_i-1)/g_0 \pmod{g_1}$. Therefore there are at most v choices for x, and so for n, modulo g_1. Therefore

$$E_4 \ll \sum_{\substack{g_0 \\ g_0 g_1 = g}} \sum_{g_1} g^{\varepsilon} E_5$$

with

$$E_5 = \sum_{l_0} (MR(j_0 f_1 \dots f_{2k} g_1)^{-1}+1) \min\left(P^{k-3}, \left(h_0^3 e_1^3 e_2^3 e_3^3 M^{2k-1} j_0 f_1 \dots f_{2k} l_0 g_0 \left|\alpha - \frac{b}{r}\right|\right)^{-1}\right)$$

and l_0 satisfying $l_0 \leqslant MR/(g_0 j_0 f_1 \dots f_{2k})$. The total contribution to E_2 from the "+1" part is

$$\ll (rP)^{2\varepsilon} H \sum_{j_0 \leqslant MR} (MR/j_0)\, P^{k-3} \ll P^{k-3} (rP)^{3\varepsilon} HMR.$$

Summing the rest over l_0 gives a contribution to E_5

$$\ll \left(\frac{MR}{j_0 f_1 \dots f_{2k}}\right)^2 \frac{P^\varepsilon}{g} \min\left(P^{k-3}, \left(h_0^3 e_1^3 e_2^3 e_3^3 M^{2k}\left|\alpha - \frac{b}{r}\right|\right)^{-1}\right).$$

Clearly $f_1 \dots f_{2k} \geq f^{1/2k}$. Hence on performing the summation over g_0, g_1 and j_0 we obtain a contribution to E_3 of

$$\ll \sum_{h_0 \leq H/(e_1e_2e_3)} (fg)^{-1/k} P^\varepsilon M^2 R^2 \min\left(P^{k-3}, \left(h_0^3 e_1^3 e_2^3 e_3^3 M^{2k}\left|\alpha - \frac{b}{r}\right|\right)^{-1}\right)$$

$$\ll (e_1e_2e_3)^{-1}(fg)^{-1/k} P^{k-3+\varepsilon} HM^2R^2 \min\left(1, \left(P^{k-3}H^3M^{2k}\left|\alpha - \frac{b}{r}\right|\right)^{-1/3}\right).$$

Obviously $e_1e_2e_3 \geq e^{1/3} \geq e^{1/k}$ where e is as above and, by (2.23),

$$P^{k-3}H^3M^{2k} = Q^k.$$

It follows, therefore, that

$$E_2 \ll E_0 + P^{k-3+\varepsilon}HMR$$

where

$$E_0 = P^{k-3+\varepsilon}HM^2R^2(r + Q^k|\alpha r - b|)^{-1/k}.$$

To summarise, we have shown that

$$E(\alpha) \ll E_0 + P^{k-3+\varepsilon}HMR^2.$$

If $r + Q^k|\alpha r - b| \geq \frac{1}{2}M^k$, then we are done, so we may suppose that

$$r + Q^k|\alpha r - b| < \frac{1}{2}M^k.$$

Therefore

$$\begin{aligned} rq\left|\frac{a}{q} - \frac{b}{r}\right| &< \frac{1}{2}qM^kQ^{-k} + rX^{-1} \\ &< \frac{1}{2}XM^kQ^{-k} + \frac{1}{2}M^kX^{-1} \\ &\leq 1. \end{aligned}$$

Hence $a=b$, $q=r$ and the lemma ensues.

When $k=4$ we require a modified argument and obtain a slightly weaker conclusion.

LEMMA 3.3. *Suppose that* $k=4$, *that* (2.24) *holds, that*

$$1\leqslant Z\leqslant \min(P^{2/3}M^{-19/6}R^{-7/6}, M^{12}P^{-2}),$$

that $Z^4\leqslant X\leqslant Q^4Z^{-4}$, *and that* $(a,q)=1$, $q\leqslant X$ *and* $|\alpha-a/q|\leqslant q^{-1}X^{-1}$. *Then*

$$E(\alpha)\ll\frac{P^{1+\varepsilon}HM^2R^2}{(q+Q^4|\alpha q-a|)^{1/4}}+P^{1+\varepsilon}HM^2R^2Z^{-1}.$$

Proof. By (3.3),

$$E(\alpha)\ll\sum_{h\leqslant H}\sum_{j\leqslant 4P}\left|\sum_{M<m\leqslant MR}e(\alpha h^3jm^8)\right|^2$$

$$\ll PHMR+\sum_{M<m_1<m_2\leqslant MR}\sum\sum_{h\leqslant H}\min(P,\|\alpha h^3(m_2^8-m_1^8)\|^{-1}).$$

We note that

$$Z\leqslant M \tag{3.5}$$

since otherwise $M<P^{2/3}M^{-19/6}$ and $M<M^{12}P^{-2}$ which leads to a contradiction. Hence

$$E(\alpha)\ll PHM^2R^2Z^{-1}+\sum_{\substack{M<m_1<m_2\leqslant MR\\ \|\alpha h^3(m_2^8-m_1^8)\|<\frac14 Z/P}}\sum\sum_{h\leqslant H}\min(P,\|\alpha h^3(m_2^8-m_1^8)\|^{-1}. \tag{3.6}$$

For a given pair m_1, m_2 with $M<m_1,m_2\leqslant MR$ choose s, c with $(c,s)=1$, $s\leqslant 2H^3$, $|\alpha(m_2^8-m_1^8)-c/s|\leqslant 1/(2sH^3)$. In addition, for h with $h\leqslant H$ and $\|\alpha h^3(m_2^8-m_1^8)\|<\frac14 Z/P$ choose b so that $|\alpha h^3(m_2^8-m_1^8)-b|<\frac14 Z/P$. Then

$$\left|\frac{b}{h^3}-\frac{c}{s}\right|h^3s<\frac14 sZP^{-1}+h^3/(2H^3)$$

$$\leqslant\frac12 H^3ZP^{-1}+\frac12$$

$$\leqslant 1$$

since $H^3ZP^{-1}=ZP^2M^{-12}\leqslant 1$. Hence $bs=ch^3$, whence $s|h^3$. Let s_3^3 denote the largest cube divisor of s, let s_2^2 denote the largest square divisor of s/s_3^3 and let $s_1=ss_2^{-2}s_3^{-3}$. Thus $s_1s_2s_3|h$. Let $h_0=h/(s_1s_2s_3)$. Then $b=ch_0^3s_1^2s_2$. Thus the multiple sum on the right of (3.6) is

$$\ll \sum_{M<m_1<m_2\leqslant MR}\sum \sum_{h_0\leqslant H/(s_1s_2s_3)} \min\left(P, (h_0s_1s_2s_3)^{-3}\left|\alpha(m_2^8-m_1^8)-\frac{c}{s}\right|^{-1}\right)$$

$$\ll \sum_{M<m_1<m_2\leqslant MR}\sum (s_1s_2s_3)^{-1}\min\left(PH, P^{2/3}\left|\alpha(m_2^8-m_1^8)-\frac{c}{s}\right|^{-1/3}\right).$$

Therefore

$$E(\alpha)\ll PHM^2R^2Z^{-1}+E_0$$

where

$$E_0=\sum_{\substack{M<m_1<m_2\leqslant MR\\ s\leqslant Z^3,\, PH^3|\alpha(m_2^8-m_1^8)s-c|<\frac{1}{32}Z^3}}\sum PH\min(s^{-1/3},(PH^3|\alpha(m_2^8-m_1^8)s-c|)^{-1/3}),$$

Put $(m_2,m_1)=j$, $n=m_1/j$, $l=(m_2-m_1)/j$. Then

$$j\leqslant MR,\quad l\leqslant MR/j,$$
$$M/j<n\leqslant MR/j,$$
$$M/j<n+l\leqslant MR/j,$$
$$(n,n+l)=1,$$

and now s and c will depend on j, l, and n.

Given $j\leqslant MR$ and $l\leqslant MR/j$, choose d, t with $(d,t)=1$, $t\leqslant 16(MR)^7Z^3$, $|\alpha j^8l-d/t|\leqslant 1/(16t(MR)^7Z^3)$, and for brevity write

$$D=((n+l)^8-n^8)/l.$$

Then

$$\left|\frac{c}{sD}-\frac{d}{t}\right| tsD < \frac{sD}{16(MR)^7 Z^3}+\frac{tZ^3}{32PH^3}$$
$$\leqslant \frac{1}{2}+\frac{1}{2}(MR)^7 Z^6 P^{-1} H^{-3}$$
$$\leqslant 1$$

since $Z^6 \leqslant P^4 M^{-19} R^{-7}$. Thus $ct=dsD$, so that $s|t$. Let $t_0=t/s$. Then $ct_0=dD$. Hence $t_0|D$. Thus $(n+l)^8 \equiv n^8 \pmod{t_0}$. Since $(n, n+l)=1$ we have $(n, t_0)=1$. Let $t_1=(l, t_0)$, $t_2=t_0/t_1$, $l_0=l/t_1$. Thus $(l_0, t_2)=1$. Since $t_0|D$ we have

$$8n^7 \equiv ((n+l_0 t_1)^8-n^8/l_0 t_1 = D \equiv 0 \pmod{t_1}.$$

Hence $t_1|8$. Moreover $(n+l_0 t_1)^8 \equiv n^8 \pmod{t_0}$. An argument allied to one in the previous lemma shows that n lies in one of $\ll t_0^{\varepsilon}$ residue classes modulo t_2. Since $s=t/t_0$ and $t_2 \gg t_0$ it follows that

$$E_0 \ll \sum_{j\leqslant MR}\sum_{l\leqslant MR/j}\sum_{t_0|t} t_0^{\varepsilon}\left(\frac{MR}{jt_0}+1\right)\left(\frac{t_0}{t}\right)^{1/3} PH \min\left(1, \left(PH^3\left(\frac{M}{j}\right)^7 \left|\alpha j^8 l-\frac{d}{t}\right|\right)^{-1/3}\right).$$

The "+1" part can be bounded trivially. Thus, by (3.5),

$$E_0 \ll P^{1+\varepsilon} HM^2R^2Z^{-1}+P^{1+\varepsilon}HMRE_1$$

where

$$E_1 = \sum_{\substack{j\leqslant MR \quad l\leqslant MR/j \\ t\leqslant Z^3 j^{-3},\, PH^3(M/j)^7 |\alpha j^8 lt-d|<\frac{1}{4}Z^3 j^{-3}}} j^{-1} t^{-1/3} \min\left(1, \left(PH^3\left(\frac{M}{j}\right)^7 \left|\alpha j^8 l-\frac{d}{t}\right|\right)^{-1/3}\right).$$

Since $t \leqslant Z^3 j^{-3}$ we have $j \leqslant Z$.

For a given j with $j\leqslant Z$, choose e, u so that $(e, u)=1$, $u \leqslant 2Z^3 MRj^{-4}$, $|\alpha j^8 - e/u| \leqslant j^4/(2uZ^3MR)$. Then

$$\left|\frac{d}{lt}-\frac{e}{u}\right| ult < \frac{ltj^4}{2Z^3MR}+\frac{uZ^3 j^{-3}}{4PH^3(M/j)^7}$$
$$\leqslant \frac{1}{2}+\frac{Z^6 MR}{2PH^3 M^7}$$
$$\leqslant 1$$

since $Z^6MRP^{-1}H^{-3}M^{-7}=Z^6P^{-4}M^6R\leqslant 1$. Therefore $du=elt$, so that $t|u$. Let $u_0=u/t$. Then $u_0|l$. Let $l_0=l/u_0$. Then

$$E_1\ll\sum_{j\leqslant Z}j^{-1}\sum_{u_0|u}\left(\frac{u_0}{u}\right)^{1/3}\min\left(\frac{MR}{ju_0},\frac{1}{u_0}\left(\frac{MR}{j}\right)^{2/3}\left(PH^3(M/j)^7\left|\alpha j^8-\frac{e}{u}\right|\right)^{-1/3}\right).$$

Therefore

$$E_1\ll P^{\varepsilon}MRZ^{-1}+P^{\varepsilon}MRE_2$$

where

$$E_2=\sum_j u^{-1/3}j^{-2}\min\left(1,\left(PH^3(M/j)^8\left|\alpha j^8-\frac{e}{u}\right|\right)^{-1/3}\right)$$

and the sum is over the j with $j\leqslant Z$, $u\leqslant Z^3/j^3$ and $PH^3(M/j)^8\ |\alpha j^8u-e|<\frac{1}{4}Z^3/j^3$.

Now choose f, v with $(f,v)=1$, $v\leqslant 2Z^8$, $|\alpha-f/v|\leqslant 1/(2vZ^8)$. Thus

$$\begin{aligned}\left|\frac{e}{j^8u}-\frac{f}{v}\right|j^8uv&<\frac{j^8u}{2Z^8}+\frac{vZ^3j^5}{4PH^3M^8}\\&\leqslant\frac{1}{2}+\frac{Z^{16}M^4}{2P^4}\\&\leqslant 1\end{aligned}$$

since $Z^{16}M^4P^{-4}\leqslant Z^6M^{14}P^{-4}\leqslant P^4M^{-19}M^{14}P^{-4}\leqslant 1$. Hence $ev=fj^8u$, so that $u|v$. Let $v_0=v/u$. Then $v_0|j^8$. Write $v_0=v_1v_2^2\dots v_8^8$ where v_8^8 is the largest eighth power dividing v_0, v_7^7 is the largest seventh power dividing v_0/v_8^8, and so on. Thus $v_1\dots v_8|j$. Let $j_0=j/(v_1\dots v_8)$. Then $v_1\dots v_8\geqslant v_0^{1/8}$, $uv_1v_2^2\dots v_8^8=v$ and $j\leqslant MR/(v_1\dots v_8)$. Hence

$$\begin{aligned}E_2&\ll\sum_u\sum_{v_1}\dots\sum_{v_8}\sum_{j_0}u^{-1/3}v_0^{-1/4}j_0^{-2}\min\left(1,\left(PH^3m^8\left|\alpha-\frac{f}{v}\right|\right)^{-1/3}\right)\\&\ll P^{\varepsilon}v^{-1/4}\min\left(1,\left(Q^4\left|\alpha-\frac{f}{v}\right|\right)^{-1/3}\right).\end{aligned}$$

Therefore, collecting together our estimates, we obtain

$$E(\alpha)\ll\frac{P^{1+\varepsilon}HM^2R^2}{(v+Q^4|\alpha v-f|)^{1/4}}+P^{1+\varepsilon}HM^2R^2Z^{-1}.$$

If $v+Q^4|\alpha v-f|\geqslant\frac{1}{2}Z^4$, then we are done, so we may suppose that

$$v+Q^4|\alpha v-f|<\frac{1}{2}Z^4.$$

Therefore

$$\begin{aligned}\left|\frac{a}{q}-\frac{f}{v}\right|qv&<\frac{1}{2}qZ^4Q^{-4}+vX^{-1}\\&<\frac{1}{2}XZ^4Q^{-4}+\frac{1}{2}Z^4X^{-1}\\&\leqslant 1.\end{aligned}$$

Hence $a=f$, $q=v$, and the lemma follows.

The case $k=3$ requires yet another variant of our argument.

LEMMA 3.4. *Suppose that* $k=3$, *that* $MR\leqslant P^{1/7}$, *that* $M^3\leqslant X\leqslant Q^3M^{-3}$, *and that* $(a,q)=1$, $q\leqslant X$ *and* $|\alpha-a/q|\leqslant q^{-1}X^{-1}$. *Then*

$$E(\alpha)\ll\frac{P^{\varepsilon}HM^2R^2}{(q+Q^3|\alpha q-a|)^{1/3}}+P^{\varepsilon}HMR^2.$$

Proof. Since $M\leqslant P^{1/7}R^{-1}$ and $H=PM^{-3}$ we have

$$H^{3/4}\leqslant HM^{-1}. \tag{3.7}$$

By (3.3),

$$E(\alpha)\ll HMR+\left|\sum_{M<m_1<m_2\leqslant MR}\sum\sum_{h\leqslant H}e(2\alpha(m_2^6-m_1^6)h^3)\right|.$$

For a given pair m_1, m_2 with $M\leqslant m_1<m_2\leqslant MR$ we choose b, r so that $(b,r)=1$, $r\leqslant 6H^2$, $|2\alpha(m_2^6-m_1^6)-b/r|\leqslant 1/(6rH^2)$. If $r>H$, then by Weyl's inequality (Lemma 2.4 of [Va2]) we have, by (3.7)

$$\sum_{h\leqslant H}e(2\alpha(m_2^6-m_1^6)\,h^3)\ll H^{\frac{3}{4}+\varepsilon}\ll P^{\varepsilon}HM^{-1}.$$

If $r \leqslant H$, then, by Theorem 4.1 and Lemma 4.6 of [Va2],

$$\sum_{h \leqslant H} e(2\alpha(m_2^6 - m_1^6) h^3 \ll r^{-1/3} \min\left(H, \left|2\alpha(m_2^6 - m_1^6) - \frac{b}{r}\right|^{-1/3}\right) + r^{\frac{1}{2}+\varepsilon}.$$

Hence

$$E(\alpha) \ll E_0 + P^\varepsilon HMR^2$$

where

$$E_0 = \sum_{\substack{M < m_1 < m_2 \leqslant MR \\ m_1, m_2 \in \mathscr{A}}} \sum r^{-1/3} \min\left(H, \left|2\alpha(m_2^6 - m_1^6) - \frac{b}{r}\right|^{-1/3}\right)$$

and $\mathscr{A}$ is the set of ordered pairs m_1, m_2 for which $r < \frac{1}{12}M^3$ and $|2\alpha(m_2^6 - m_1^6)r - b| < \frac{1}{2}M^3H^{-3}$.

Given such a pair m_1, m_2 put $j = (m_1, m_2)$, $n = m_1/j$, $l = (m_2 - m_1)/j$. Thus

$$E_0 \leqslant \sum_j \sum_l \sum_n r^{-1/3} \min\left(H, \left|2\alpha j^6 lD - \frac{b}{r}\right|^{-1/3}\right)$$

where j, l and n satisfy

$$j \leqslant MR,$$

$$l \leqslant MR/j,$$

$$(n, n+l) = 1,$$

$$M/j < n \leqslant MR/j,$$

$$M/j < n + l \leqslant MR/j,$$

$$jn,\ jn + jl \in \mathscr{A}.$$

and we have written D for $((n+l)^6 - n^6)/l$.

Given $j \leqslant MR$, $l \leqslant MR/j$, choose c, s so that $(c, s) = 1$, $s \leqslant H^3M^{-3}$ and $|2\alpha j^6 l - c/s| \leqslant s^{-1}M^3H^{-3}$. Thus, for any n in the innermost sum we have

$$\left|\frac{c}{s}-\frac{b}{rD}\right| sDr < DrM^3H^{-3}+\frac{1}{2}sM^3H^{-3}$$
$$\leqslant \frac{6}{12}\left(\frac{MR}{j}\right)^5 M^6h^{-3}+\frac{1}{2}$$
$$\leqslant 1$$

since $M \leqslant P^{1/7}R^{-1} \leqslant P^{3/20}R^{-1/4}$. Hence $crD=bs$, so that $r|s$. Let $s_0=s/r$. Then $s_0|D$. Hence the innermost sum in E_0 is

$$\ll \sum_{s_0|s}\left(\frac{s_0}{s}\right)^{1/3}\sum_n \min\left(H, \left(\left|2aj^6l-\frac{c}{s}\right|(M/j)^5\right)^{-1/3}\right)$$

where the sum over n is now over n with $n \leqslant MR/j$, $(n, n+l)=1$ and $((n+l)^6-n^6)/l \equiv 0 \pmod{s_0}$. Now, much as in the proof of the previous lemma we find that the number of such n is

$$\ll \left(\frac{MR}{js_0}+1\right)s_0^{\varepsilon}.$$

Therefore

$$E(\alpha) \ll P^{\varepsilon}HMR^2+P^{\varepsilon}MRE_1$$

where

$$E_1=\sum_{j\leqslant MR}\sum_{l\in\mathscr{L}} s^{-1/3}j^{-1}\min\left(H, \left(\left|2aj^6l-\frac{c}{s}\right|\left(\frac{M}{j}\right)^5\right)^{-1/3}\right)$$

and $\mathscr{L}$ is the set of l for which $l \leqslant MR/j$, $s<\frac{1}{4}M^3$ and

$$|2aj^6ls-c|(M/j)^5<\frac{1}{2}M^3H^{-3}.$$

Now given $j \leqslant MR$ choose d, t so that $(d,t)=1$, $t \leqslant M^4Rj^{-1}$ and

$$|\alpha j^6-d/t| \leqslant j/(tM^4R).$$

Then

$$\left|\frac{c}{2sl}-\frac{d}{t}\right|2slt<\frac{tM^3H^{-3}}{2(M/j)^5}+\frac{2slj}{M^4R}$$
$$<\frac{M^6R^5}{2H^3}+\frac{1}{2}$$
$$\leqslant 1$$

since $M\leqslant P^{1/7}R^{-1}\leqslant P^{1/5}R^{-1/3}$. Therefore $ct=2dsl$ so that $s|t$. Let $t_0=t/s$. Then $t_0|2l$. Put $l_0=2l/t_0$. Then

$$E_1\ll\sum_{j\leqslant MR}\sum_{t_0|t}\left(\frac{t_0}{t}\right)^{1/3}\sum_{l_0\leqslant 2MR/(jt_0)}j^{-1}\min\left(H,\left(\left|\alpha j^6-\frac{d}{t}\right|l_0t_0\left(\frac{M}{j}\right)^5\right)^{-1/3}\right)$$
$$\ll\sum_{j\leqslant MR}P^{\varepsilon}MRt^{-1/3}j^{-2}\min\left(H,\left(\left|\alpha j^6-\frac{d}{t}\right|\left(\frac{M}{j}\right)^6\right)^{-1/3}\right).$$

Therefore

$$E(\alpha)\ll P^{\varepsilon}HMR^2+P^{\varepsilon}M^2R^2E_2$$

where

$$E_2=\sum_{j\in J}t^{-1/3}j^{-2}\min\left(H,\left|\alpha j^6-\frac{d}{t}\right|\left(\frac{M}{j}\right)^6\right)^{-1/3}$$

and $\mathscr{J}$ is the set of $j\leqslant M$ for which $t<\frac{1}{2}M^3$ and

$$|\alpha j^6t-d|\,(M/j)^6<\tfrac{1}{2}M^3H^{-3}.$$

Now choose e, u so that $(e,u)=1$, $u\leqslant M^9$, $|\alpha-e/u|\leqslant u^{-1}M^{-9}$. Thus for $j\in\mathscr{J}$,

$$\left|\frac{e}{u}-\frac{d}{j^6t}\right|uj^6t<\frac{uM^3H^{-3}}{2(M/j)^6}+j^6tM^{-9}$$
$$<\frac{1}{2}M^{12}H^{-3}+\frac{1}{2}$$
$$\leqslant 1$$

since $M\leqslant P^{1/7}R^{-1}$. Therefore $etj^6=du$, whence $t|u$. Let $u_0=u/t$. Then $u_0|j^6$. Let u_6^6 denote the largest sixth power dividing u_0, u_5^5 the largest fifth power dividing u_0/u_6^6, and so on.

Thus $u_0=u_1u_2^2\ldots u_6^6$, $u_1u_2\ldots u_6|j$ and $u_1u_2\ldots u_6\geqslant u_0^{1/6}$. Therefore

$$E_2\ll\sum_{u_0|u}(u_0/u)^{1/3}\sum_{j_0\leqslant M}j_0^{-2}u_0^{-1/3}\min\left(H,\left(\left|\alpha-\frac{e}{u}\right|M^6\right)^{-1/3}\right)$$

$$\ll P^\varepsilon H(u+Q^3|\alpha u-e|)^{-1/3}.$$

Hence

$$E(\alpha)\ll\frac{P^\varepsilon HM^2R^2}{(u+Q^3|\alpha u-e|)^{1/3}}+P^\varepsilon HMR^2.$$

The proof of the lemma can now be concluded in the same manner as that of Lemma 3.2.

The previous four lemmas are of greatest utility on minor arcs. Whilst they do give some information on major arcs it is important to establish a more precise estimate. We do this in the next lemma.

LEMMA 3.5. *Suppose that* $(a,q)=1$, $\beta=\alpha-a/q$, *and*

$$k(k-1)\,3^kqP^{k-2}HR^{k(k-2)}|\beta|\leqslant 1. \tag{3.8}$$

Then

$$F_1(\alpha)\ll\frac{PHMRq^\varepsilon}{(q+Q^k|\alpha q-a|)^{1/(k-1)}}+HMRq^{\frac{k-2}{k-1}+\varepsilon}. \tag{3.9}$$

Proof. By (2.35) and (2.31),

$$F_1(\alpha)=\sum_{h\leqslant H}\sum_{M<m\leqslant MR}S(\alpha,h,m) \tag{3.10}$$

where

$$S(\alpha,h,m)=\sum_{z\leqslant 2P}e(\alpha m^{-k}(z+hm^k)^k-\alpha m^{-k}(z-hm^k)^k). \tag{3.11}$$

Hence, on writing $\alpha=a/q+\beta$, sorting the terms in $S(\alpha,h,m)$ according to the residue class r of z modulo q, and using the fact that

$$\sum_{-\frac{1}{2}q<b\leqslant\frac{1}{2}q}e(b(r-z)/q)$$

is q or 0 according as $r\equiv z \pmod q$ or $r\not\equiv z \pmod q$, we obtain

$$S(\alpha,h,m)=q^{-1}\sum_{-\frac{1}{2}q<b\le\frac{1}{2}q}\sigma(q,a,b,h,m)\,T(\beta,b,h,m) \tag{3.12}$$

where

$$\sigma(q,a,b,h,m)=\sum_{r=1}^{q}e\left(\frac{a}{q}m^{-k}(r+hm^k)^k-\frac{a}{q}m^{-k}(r-hm^k)^k+\frac{b}{q}r\right)$$

and

$$T(\beta,b,h,m)=\sum_{z\le 2P}e\left(\beta m^{-k}(z+hm^k)^k-\beta m^{-k}(z-hm^k)^k-\frac{b}{q}z\right).$$

When k is even, let d denote the greatest common division of

$$q,2akh,2a\binom{k}{3}h^3m^{2k},\dots,2a\binom{k}{k-3}h^{k-3}m^{k(k-4)},2akh^{k-1}m^{k(k-2)}+b$$

and when k is odd, let d denote the greatest common divisor of

$$q,2akh,2a\binom{k}{3}h^3m^{2k},\dots,2a\binom{k}{k-2}h^{k-2}m^{k(k-3)},\ b.$$

Then by Theorem 7.1 of [Va2],

$$\sigma(q,a,b,h,m)\ll d(q/d)^{\frac{k-2}{k-1}+\varepsilon}.$$

If k is odd, then $d\ll(q,h,b)$, and if k is even, then $d\ll(q,h,2akh^{k-1}m^{k(k-2)}+b)=(q,h,b)$. Thus

$$\sigma(q,a,b,h,m)\ll q^{\frac{k-2}{k-1}+\varepsilon}(q,h,b)^{\frac{1}{k-1}}. \tag{3.13}$$

Let

$$\phi(\gamma)=\beta m^{-k}(\gamma+hm^k)^k-\beta m^{-k}(\gamma-hm^k)^k-\frac{b}{q}\gamma. \tag{3.14}$$

Then

$$\frac{b}{q}+\phi'(\gamma)=k(k-1)\beta m^{-k}\int_{\gamma-hm^k}^{\gamma+hm^k}\psi^{k-2}d\psi,$$

so that when $|\gamma|\leqslant 2P$ we have

$$\left|\frac{b}{q}+\phi'(\gamma)\right|\leqslant 2k(k-1)|\beta|\,h(2P+hm^k)^{k-2}\leqslant\frac{2}{9q}.$$

Thus, when $-\frac{1}{2}q<b\leqslant\frac{1}{2}q$ and $|\gamma|\leqslant 2P$ we have

$$|\phi'(\gamma)|\leqslant\frac{1}{2}+\frac{2}{9}<\frac{3}{4}$$

and if moreover $b\neq 0$, then

$$|\phi'(\gamma)|>\frac{|b|}{2q}.$$

Therefore, by Lemma 4.2 of [Va2], we have

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

where

$$I(\beta,b,h,m,u)=\int_0^{2P}e(\phi(\gamma)-\gamma u)d\gamma. \tag{3.15}$$

It follows by integration by parts that

$$I(\beta,b,h,m,\pm 1)\ll 1$$

and, when $b\neq 0$, that

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

Therefore

$$T(\beta,0,h,m)=I(\beta,0,h,m,0)+O(1)$$

and, when $b\neq 0$,

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

Hence, by (3.12) and (3.13),

$$S(\alpha,h,m)=q^{-1}\sigma(q,a,0,h,m)I(\beta,0,h,m,0)+O\Bigg(\sum_{1\le b\le\frac{1}{2}q}|b|^{-1}q^{\frac{k-2}{k-1}+\varepsilon}(q,b)^{\frac{1}{k-1}}\Bigg)$$
$$\ll q^{\varepsilon-\frac{1}{k-1}}(q,h)^{\frac{1}{k-1}}|I(\beta,0,h,m,0)|+q^{\frac{k-2}{k-1}+2\varepsilon}.$$

By (3.14), (3.15) and Theorem 7.3 of [Va2], we have

$$I(\beta,0,h,m,0)\ll P(1+|\beta|\,hP^{k-1})^{-1/(k-1)}.$$

Thus

$$S(\alpha,h,m)\ll q^{\varepsilon-\frac{1}{k-1}}(q,h)^{\frac{1}{k-1}}P(1+|\beta|\,hP^{k-1})^{-1/(k-1)}+q^{\frac{k-2}{k-1}+\varepsilon}.$$

Therefore, by (3.10),

$$F_1(\alpha)\ll MRq^{\varepsilon-\frac{1}{k-1}}\sum_{h\le H}(q,h)^{\frac{1}{k-1}}\min(P,(|\beta|\,h)^{-\frac{1}{k-1}})+HMRq^{\frac{k-2}{k-1}+\varepsilon}.$$

On writing $d=(q,h)$ and recalling that $P^{k-1}H=Q^k$ we obtain

$$\sum_{h\le H}(q,h)^{\frac{1}{k-1}}\min(P,(|\beta|\,h)^{-\frac{1}{k-1}})\ll\sum_{d|q}d^{\frac{1}{k-1}}\min(PHd^{-1},H^{\frac{k-2}{k-1}}|\beta|^{-\frac{1}{k-1}}d^{-1})$$
$$\ll q^{\varepsilon}PH\min(1,(Q^k|\beta|)^{-\frac{1}{k-1}}).$$

The lemma now follows.

Having established suitable estimates for the underlying exponential sums we are now in a position to establish a relationship between T_s, and S_{s-1} and S_s that is particularly valuable when $S_s\approx P^\lambda$ with λ close to $2s-k$.

LEMMA 3.6. *Suppose that* $k\ge4$, $s\ge k-1$ *and* (2.22), (2.23) and (2.24) *hold. Then*

$$T_s(P,R,\theta)\ll(PMR+PHMR^4(ZP)^{-2^{2-k}})S_{s-1}(Q,R)+P^{1+\sigma+\varepsilon}HMRQ^{-k/s}S_s(Q,R)^{1-1/s}$$

where

$$Z=\begin{cases}\min(M^{12}P^{-2},P^{2/3}R^{-7/6}M^{-19/6}) & (k=4),\\ M & (k\ge5),\end{cases}$$

$$\sigma = \begin{cases} 0 & (s \geqslant 2k-2), \\ \dfrac{2}{s} - \dfrac{1}{k-1} & (2k-2 > s \geqslant k-1). \end{cases}$$

Proof. Let $\mathfrak{m}$ denote the set of points in $[0, 1]$ with the property that whenever there are a, q with $(a, q)=1$ and

$$k(k-1)\,3^k q P^{k-2} H R^{k(k-2)} \left|\alpha - \frac{a}{q}\right| \leqslant 1, \tag{3.16}$$

then $q>P$. Further, let

$$\mathfrak{M} = [0, 1] \setminus \mathfrak{m}.$$

First of all suppose that $\alpha \in \mathfrak{m}$. Choose b, r with

$$(b, r) = 1, \quad r \leqslant P^{k-2}H \quad \text{and} \quad |\alpha r - b|\, P^{k-2}H \leqslant 1.$$

Then, by Lemma 3.1,

$$D(\alpha) \ll \frac{P^{k-1+\varepsilon}}{r + Q^k|\alpha r - b|} + HP^{k-2+\varepsilon}.$$

Since $\alpha \in \mathfrak{m}$, either $r>P$ or $Q^k|\alpha r - a| \gg PR^{-k(k-2)}$. Thus

$$D(\alpha) \ll P^{k-2} H R^{k(k-2)}. \tag{3.17}$$

Clearly when $k \geqslant 4$ we have

$$Z^k \leqslant P^{k-2} H \leqslant Q^k Z^{-k}$$

(recall that when $k=4$ we have (3.5)). Hence, by Lemmas 3.2 and 3.3,

$$E(\alpha) \ll \frac{P^{k-3} H M^2 R^2}{(r + Q^k|\alpha r - b|)^{1/k}} + P^{k-3+\varepsilon} H M^2 R^2 Z^{-1} \ll P^{k-3+\varepsilon} H M^2 R^k Z^{-1}.$$

Hence, by (3.1), (3.17) and (2.37),

$$F_1(\alpha) \ll P^{1+\varepsilon} H M R^{1+k(k-1)2^{2-k}} (PZ)^{-2^{2-k}},$$

whence

$$F_1(\alpha) \ll P^{1+\varepsilon} H M R^4 (PZ)^{-2^{2-k}} \quad (\alpha \in \mathfrak{m}). \tag{3.18}$$

Now suppose that $\alpha \in \mathfrak{M}$. We note that, by Dirichlet's theorem on diophantine approximation, there are q, a with $(a,q)=1$ and satisfying (3.16). Moreover since α is not in $\mathfrak{m}$, there are a, q with $(a,q)=1$, $q\leqslant P$ and satisfying (3.16). Furthermore, as $0\leqslant\alpha\leqslant1$ we have $0\leqslant a\leqslant q$. Thus, by Lemma 3.5,

$$F_1(\alpha) \ll \frac{P^{1+\varepsilon}HMR}{(q+Q^k|\alpha q-a|)^{1/(k-1)}} + P^{\frac{k-2}{k-1}+\varepsilon}HMR. \tag{3.19}$$

Let $\mathfrak{M}(q,a)$ denote the set of α in $[0,1]$ for which (3.16) holds. Note that the $\mathfrak{M}(q,a)$ with $0\leqslant a\leqslant q\leqslant P$ are disjoint. We now define $F^*(\alpha)$ on $[0,1]$ by taking $F^*(\alpha)$ to be 0 when $\alpha\in\mathfrak{m}$ and to be

$$\frac{P^{1+\varepsilon}HMR}{(q+Q^k|\alpha q-a|)^{1/(k-1)}}$$

when $\alpha\in\mathfrak{M}(q,a)$ with $0\leqslant a\leqslant q\leqslant P$.

As $Z\leqslant M\leqslant P^{1/k}$ and $k\geqslant4$ we have

$$(PZ)^{2^{2-k}}\leqslant P^{1/(k-1)}.$$

Therefore, by (3.18), (3.19) and (2.36)

$$T_s(P,R,\theta)\ll(PMR+P^{1+\varepsilon}HMR^4(PZ)^{-2^{2-k}})\,S_{s-1}(Q,R)+I$$

where

$$I=\int_{\mathfrak{M}}F^*(\alpha)\,|f(2^k\alpha;Q,R)|^{2s-2}\,d\alpha.$$

By Hölder's inequality,

$$I\ll J^{1/s}S_s(Q,R)^{1-1/s}$$

where

$$J=\int_{\mathfrak{M}}F^*(\alpha)^s\,d\alpha.$$

A straightforward calculation shows that

$$J\ll(P^{1+\varepsilon}HMR)^sP^{\varepsilon}Q^{-k}\sum_{q\leqslant P}q^{1-s/(k-1)}$$

and the lemma follows.

When $k=3$ we can obtain a more precise result. In principle such a result could be obtained for larger k but it would be valid only when $s>2^{k-2}$, which is too large to be useful when $k\geq 4$.

LEMMA 3.7. *Suppose that* $k=3$, $MR\leq P^{1/7}$ *and* (2.22) *and* (2.23) *hold. Then*

$$T_3(P,R,\theta)\ll P^{3+\varepsilon}M^{-1}R+P^{\frac{7}{6}+\varepsilon}M^{-\frac{3}{2}}RS_3(Q,R)^{\frac{2}{3}}.$$

Proof. Let $\mathfrak{m}$ denote the set of points α in $[0,1]$ with the property that whenever there are a, q with $(a,q)=1$ and

$$PH|\alpha q-a|\leq 1,$$

then $q<P$, and let

$$\mathfrak{M}=[0,1]\setminus\mathfrak{m}.$$

Let $\alpha\in\mathfrak{m}$ and choose a, q so that $(a,q)=1$, $|\alpha q-a|\leq H^{-1}P^{-1}$ and $q\leq PH$. Then $q>P$. Hence, by (3.1) and Lemmas 3.1 and 3.4 we have

$$F_1(\alpha)\ll P^{\varepsilon}(PH)^{1/2}(HMR^2)^{1/2}=P^{\varepsilon}HR(PM)^{1/2}.$$

By (2.32), (2.34), (2.35) and Lemma 2.3 with $j=1$ we have

$$\int_0^1|F_1(\alpha)|^2\,d\alpha\ll P^{1+\varepsilon}HMR.$$

Thus

$$\int_{\mathfrak{m}}|F_1(\alpha)|^3\,d\alpha\ll P^{\varepsilon}H^2R^2(PM)^{3/2}. \tag{3.20}$$

Now suppose that $\alpha\in\mathfrak{M}$. Then α is in an interval of the form

$$\mathfrak{M}(q,a)=\{\alpha\colon|\alpha q-a|\leq H^{-1}P^{-1}\}$$

with $(a,q)=1$, $0\leq a\leq q\leq P$. Hence, by Lemmas 3.1 and 3.4, we have

$$F_1(\alpha)\ll P^{\varepsilon}\left(\frac{HP^2}{q+Q^3|\alpha q-a|}\right)^{1/2}\left(\frac{HM^2R^2}{(q+Q^3|\alpha q-a|)^{1/3}}+HMR^2\right)^{1/2}.$$

Thus

$$\int_{\mathfrak{M}(q,a)} |F_1(\alpha)|^3 d\alpha \ll \int_0^\infty \left(\frac{P^{3+\varepsilon}H^3M^3R^3}{(q+Q^3q\beta)^2} + \frac{P^{3+\varepsilon}H^3M^{3/2}R^3}{(q+Q^3q\beta)^{3/2}} \right) d\beta$$

$$\ll P^{3+\varepsilon}H^3R^3Q^{-3}(M^3q^{-2}+M^{3/2}q^{-3/2}).$$

Therefore

$$\int_{\mathfrak{M}} |F_1(\alpha)|^3 d\alpha \ll P^{3+\varepsilon}H^3M^3R^3Q^{-3}(1+P^{1/2}M^{-3/2}).$$

Hence, by (3.20) and (2.23),

$$\int_0^1 |F_1(\alpha)|^3 d\alpha \ll P^{\frac{7}{2}+\varepsilon}M^{-\frac{9}{2}}R^3.$$

Now, by (2.36) and Hölder's inequality,

$$T_3(P,R,\theta) \ll PMRS_2(Q,R) + P^{\frac{7}{6}+\varepsilon}M^{-\frac{3}{2}}RS_3(Q,R)^{\frac{2}{3}}.$$

The lemma is a consequence of this and the classical estimate

$$S_2(Q,R) \ll Q^{2+\varepsilon}$$

(see, for example, Lemma 2.5 of [Va2]).

4. Bounds for the number of solutions of the auxiliary equation

We now investigate the consequences of the reduction relations contained in Lemmas 2.1, 2.2 and 3.6. The aim is to establish bounds of the form

$$S_s(P,R) \ll P^{\lambda_s+\varepsilon}$$

when R is no larger than a small power of P. The reduction relations contained in Lemmas 2.1, 2.2 and 3.6 can be interpreted as inequalities between the permissible choices for λ_s. It is useful, therefore, to summarise below the corresponding inequalities.

(j1) *For some j with $2^j \geqslant t$ and $1 \leqslant j \leqslant k-1$ there is a θ satisfying $0 < \theta \leqslant 1/k$ such that*

$$\lambda_t \geqslant (2t-2)\theta + 1 + \lambda_{t-1}(1-\theta), \qquad (j1.1)$$

$$\lambda_t \geqslant (2t-2-k)\theta+2-2^{1-j}+\lambda_{t-1}(1-\theta), \tag{j1.2}$$

$$\lambda_t \geqslant \frac{(2t-2-k+k2^{-j})\theta+2-(j+1)2^{-j}+\lambda_{t-1}(1-\theta)(1-t2^{-j})}{1-(1-\theta)(t-1)2^{-j}}. \tag{j1.3}$$

($j2$) *For some j with $2^j \geqslant t$ and with*

(i) $j=k-2$, *or*

(ii) $j=k-4$, *or*

(iii) $1 \leqslant j \leqslant k-3$ *and $k-j$ odd,*

there is a θ satisfying $0<\theta\leqslant 1/k$ such that

$$\lambda_t \geqslant (2t-2)\theta+1+\lambda_{t-1}(1-\theta), \tag{j2.1}$$

$$\lambda_t \geqslant (2t-2-k)\theta+2-2^{1-j}+\lambda_{t-1}(1-\theta), \tag{j2.2}$$

$$\lambda_t \geqslant \frac{(2t-2-k+(k-1)2^{-j})\theta+2-(j+1)2^{-j}+\lambda_{t-1}(1-\theta)(1-t2^{-j})}{1-(1-\theta)(t-1)2^{-j}}. \tag{j2.3}$$

($k-2$) *We have $t \geqslant k-1$ and there is a θ satisfying $0<\theta\leqslant 1/k$ such that*

$$\lambda_t \geqslant (2t-2)\theta+1+\lambda_{t-1}(1-\theta), \tag{k-2.1}$$

$$\lambda_t \geqslant (2t-2-k-2^{2-k})\theta+2-2^{2-k}+\lambda_{t-1}(1-\theta), \tag{k-2.2}$$

$$\lambda_t \geqslant 2t-k+\sigma \bigg/ \left(\frac{1}{t}+\theta-\frac{\theta}{t}\right) \tag{k-2.4}$$

where

$$\sigma = \begin{cases} 0 & \text{when } t \geqslant 2k-2, \\ \dfrac{2}{t}-\dfrac{1}{k-1} & \text{when } 2k-2 > t \geqslant k-1. \end{cases} \tag{k-2.4}$$

THEOREM 4.1. *Suppose that*

$$k \geqslant 5, \quad \lambda_1 = 1, \quad \lambda_2 = 2$$

and that for each $t=3,4,\ldots,s$ at least one of ($j1$), ($j2$) or ($k-2$) holds. Then there are positive real numbers $D_1,\ldots,D_s$ such that for each positive number ε there are real numbers $C_1(\varepsilon),\ldots,C_s(\varepsilon)$ such that whenever $P \geqslant R$ we have

$$S_t(P,R) \leqslant (C_t(\varepsilon)R^{D_t})^{\log \frac{2\log P}{\log R}} P^{\lambda_t+\varepsilon}$$

for $t=1,2,\dots,s$.

Proof. This is by induction on s. The cases $s=1,2$ are classical. We may suppose, therefore, that $s\geqslant 3$ and that the case $s-1$ holds. Then, on hypothesis, there is a θ provided by ($j1$), ($j2$) or ($k-2$) with $t=s$.

We now prove by (sub)induction on n that for suitable D_s^*, $C_s^*(\varepsilon)$ we have

$$S_s(P,R)\leqslant (C_s^*(\varepsilon)R^{D_s^*})^n P^{\lambda_s+\varepsilon} \tag{4.1}$$

whenever $R^{(1-\theta)^{1-n}}\leqslant P<R^{(1-\theta)^{-n}}$.

We observe that if $\lambda_s\geqslant 2s$, then the conclusion is trivial. Moreover it follows easily by induction on t from ($j1.1$), ($j2.1$), ($k-2.1$) that $\lambda_s>s$. Hence we may assume that

$$s<\lambda_s<2s. \tag{4.2}$$

The bound (4.1) is trivial for $n\leqslant n_0(k)$ provided we take

$$D_s^*\geqslant 2s(1-\theta)^{-n_0(k)}. \tag{4.3}$$

Therefore we may suppose that

$$n>n_0(k)$$

and that (4.1) holds for n replaced by $n-1$.

Clearly it suffices to establish (4.1) when

$$0<\varepsilon<\varepsilon_0(s).$$

By Lemma 2.1 with $D=P^\theta$ we have

$$\begin{aligned} S_s(P,R)\ll &\left(\sum_{d>P^\theta}(C_s^*(\varepsilon)R^{D_s^*})^{\frac{n-1}{s}}\left(\frac{P}{d}\right)^{(\lambda_s+\varepsilon)/s}\right)^s+P^{2s(2\theta-\theta^2)}\\ &+P^{\varepsilon^2}\left(\sum_{d\leqslant P^\theta}\left(\left(\frac{P}{d}\right)^\theta R\right)^{\frac{2s-3}{s}}T_s(P/d,R,\theta)^{1/s}\right)^s. \end{aligned} \tag{4.4}$$

Suppose first of all that θ is provided by ($j1$) or ($j2$). For brevity we write

$$X=P/d,\quad J=2^{-j}.$$

By Lemma 2.2,

$$T_s(X,R,\theta) \ll (X^{1+\theta}+X^{2-2J+(1-k)\theta})RS_{s-1}(X^{1-\theta},R)$$
$$+X^{2-2(j+1)J+(1-2J)(1-k)\theta}R^{1-2J}N_j^J S_{s-1}(X^{1-\theta},R)^{1-sJ}S_s(X^{1-\theta},R)^{(s-1)J}.$$

Let $\nu=2$ when θ is provided by ($j1$) and let $\nu=1$ when θ is provided by ($j2$). In the latter case $j=k-2$, or $j=k-4$, or $1\leqslant j\leqslant k-3$ and $k-j$ is odd. Thus, by Lemma 2.3, the case $n-1$ of (4.1) and the case $s-1$ of the theorem we have

$$T_s(X,R,\theta) \ll (X^{1+\theta}+X^{2-2J+(1-k)\theta})R(C_{s-1}(\varepsilon)R^{D_{s-1}})^{\log\frac{2\log P}{\log R}}X^{(\lambda_{s-1}+\varepsilon)(1-\theta)}$$
$$+X^{\mu}R((C_{s-1}(\varepsilon)R^{D_{s-1}})^{\log\frac{2\log P}{\log R}}X^{(\lambda_{s-1}+\varepsilon)(1-\theta)})^{1-sJ}((C_s^*(\varepsilon)R^{D_s^*})^{n-1}X^{(\lambda_s+\varepsilon)(1-\theta)})^{(s-1)J}$$

where

$$\mu=2-(j+1)J+(1-k)\theta+(k-2+\nu)J\theta+\varepsilon^2.$$

Hence, by ($j1$) and ($j2$),

$$X^{\theta(2s-3)}T_s(X,R,\theta) \ll C_{s-1}(\varepsilon)^{1+n\log\frac{1}{1-\theta}}R^{\mu_1}X^{\lambda_s+\varepsilon(1-\theta)}$$
$$+C_{s-1}(\varepsilon)^{(1+n\log\frac{1}{1-\theta})(1-sJ)}C_s^*(\varepsilon)^{(n-1)(s-1)J}R^{\mu_2}X^{\mu_3+\varepsilon^2+\varepsilon(1-\theta)(1-J)}$$

where

$$\mu_1 = 1+D_{s-1}+D_{s-1}\,n\log\frac{1}{1-\theta},$$

$$\mu_2 = 1+\left(D_{s-1}+D_{s-1}\,n\log\frac{1}{1-\theta}\right)(1-sJ)+D_s^*(n-1)(s-1)J,$$

$$\mu_3 = 2-(j+1)J+(2s-2-k+(k-2+\nu)J)\theta+\lambda_{s-1}(1-\theta)(1-sJ)+\lambda_s(1-\theta)(s-1)J.$$

By ($j1.3$) and ($j2.3$),

$$\mu_3 \leqslant \lambda_s.$$

Therefore

$$X^{\theta(2s-3)}T_s(X,R,\theta) \ll \left(C_{s-1}(\varepsilon)^{n\log\frac{1}{1-\theta}}R^{\mu_1}+C_{s-1}(\varepsilon)^{n(1-sJ)\log\frac{1}{1-\theta}}C_s^*(\varepsilon)^{(n-1)(s-1)J}R^{\mu_2}\right)X^{\lambda_s+\varepsilon(1-\theta)}.$$

Hence, by (4.2),

$$\left(\sum_{d\leqslant P^{\theta}}\left(\left(\frac{P}{d}\right)^{\theta}R\right)^{\frac{2s-3}{s}}T_s(P/d,R,\theta)^{1/s}\right)^s$$
$$\ll\left(C_{s-1}(\varepsilon)^{n\log\frac{1}{1-\theta}}R^{\mu_1}+C_{s-1}(\varepsilon)^{n(1-sJ)\log\frac{1}{1-\theta}}C_s^*(\varepsilon)^{(n-1)(s-1)J}R^{\mu_2}\right)R^{2s-3}P^{\lambda_s+\varepsilon(1-\theta)}.$$

Thus, by (4.2) and (4.4), and provided that $C_s^*(\varepsilon)$ and D_s^* are large enough in terms of the implicit constant and $C_{s-1}(\varepsilon)$, and D_{s-1}, θ, s respectively, we have (4.1) as required.

Now suppose that θ is provided by $(k-2)$. Then we proceed as above but use Lemma 3.6 in place of Lemmas 2.2 and 2.3. Note that since $n>n_0(k)$ the hypotheses of Lemma 3.6 are satisfied. We thereby obtain

$$X^{\theta(2s-3)}T_s(X,R,\theta)$$
$$\ll\left(C_{s-1}(\varepsilon)^{n\log\frac{1}{1-\theta}}R^{4+D_{s-1}+D_{s-1}n\log\frac{1}{1-\theta}}+C_s^*(\varepsilon)^{(n-1)(1-\frac{1}{s})}R^{1+D_s^*(n-1)(1-\frac{1}{s})}\right)X^{\lambda_s+\varepsilon(1-\theta)}.$$

Therefore (4.1) follows from (4.4) as before.

The theorem is immediate from (4.1).

We now state a cleaner version of the above theorem.

THEOREM 4.2. *Suppose that $k\geqslant 5$, $\lambda_1=1$, $\lambda_2=2$, and that for each $t=3,4,\dots,s$ at least one of $(j1)$, $(j2)$ or $(k-2)$ holds. Suppose further that $\lambda>\lambda_s$. Then, provided that $0<\eta<\eta_0(\lambda-\lambda_s)$ and $P>P_0(\eta,s)$ we have*

$$S_s(P,P^{\eta})<P^{\lambda}.$$

Proof. In Theorem 4.1 take

$$\varepsilon=\frac{1}{2}(\lambda-\lambda_s)$$

and choose η_0 so small that whenever $0<\eta<\eta_0$ one has

$$\eta D_s\log\frac{e}{\eta}<\frac{1}{4}(\lambda-\lambda_s).$$

This gives Theorem 4.2.

In Table 4.1 are listed the optimal values of λ_s for those s with $\lambda_s>2s-k$ that arise in Theorems 4.1 and 4.2 when $5\leqslant k\leqslant 8$. Also listed are the corresponding values of θ. Moreover in the column headed by j, $(k-2)$ indicates that $(k-2)$ gives the optimal value, and otherwise $(j1)$ is satisfied with the indicated value of j unless $k-j$ is odd or $j=k-2$

k	s	j	θ	λ_s	k	s	j	θ	λ_s
4	3	2	0.1250000000	3.250000	5	3	2	0.1000000000	3.200000
	4	2	0.1338305414	4.618034		4	2	0.1000000000	4.480000
	5	$(k-2)$	0.1818181819	6.232937		5	3	0.1500000000	6.008000
						6	3	0.1635321106	7.660821
						7	$(k-2)$	0.1707317074	9.401656
						8	$(k-2)$	0.1707317074	11.186739
						9	$(k-2)$	0.1707317074	13.008516
6	3	2	0.0833333334	3.166667	7	3	2	0.0714285715	3.142858
	4	2	0.0833333334	4.402778		4	2	0.0714285715	4.346939
	5	3	0.1250000000	5.852431		5	3	0.1071428572	5.738339
	6	3	0.1305160302	7.393755		6	3	0.1078205829	7.197834
	7	3	0.1449861282	9.061597		7	3	0.1195745911	8.772051
	8	4	0.1500798248	10.802752		8	4	0.1276289272	10.439288
	9	4	0.1545352505	12.605910		9	4	0.1318261785	12.172336
	10	$(k-2)$	0.1546391753	14.440048		10	5	0.1347204930	13.957442
	11	$(k-2)$	0.1546391753	16.299834		11	5	0.1361941573	15.780403
	12	$(k-2)$	0.1546391753	18.181303		12	5	0.1375951109	17.636189
	13	$(k-2)$	0.1546391753	20.081102		13	$(k-2)$	0.1377777778	19.512981
						14	$(k-2)$	0.1377777778	21.406748
						15	$(k-2)$	0.1377777778	23.315152
						16	$(k-2)$	0.1377777778	25.236175
						17	$(k-2)$	0.1377777778	27.168080
						18	$(k-2)$	0.1377777778	29.109367
						19	$(k-2)$	0.1377777778	31.058743
						20	$(k-2)$	0.1377777778	33.015094
8	3	2	0.0625000000	3.125000	8	16	$(k-2)$	0.1228070176	24.592514
	4	2	0.0625000000	4.304688		17	$(k-2)$	0.1228070176	26.502205
	5	3	0.0937500000	5.651124		18	$(k-2)$	0.1228070176	28.422987
	6	3	0.0937500000	7.058831		19	$(k-2)$	0.1228070176	30.353498
	7	3	0.1004741178	8.555290		20	$(k-2)$	0.1228070176	32.292542
	8	4	0.1104129144	10.156457		21	$(k-2)$	0.1228070176	34.239072
	9	4	0.1142072185	11.823832		22	$(k-2)$	0.1228070176	36.192168
	10	5	0.1175704127	13.549966		23	$(k-2)$	0.1228070176	38.151025
	11	5	0.1190184991	15.317640		24	$(k-2)$	0.1228070176	40.114934
	12	5	0.1204379548	17.122450		25	$(k-2)$	0.1228070176	42.083276
	13	6	0.1213892533	18.957310		26	$(k-2)$	0.1228070176	44.055505
	14	6	0.1219220239	20.815969		27	$(k-2)$	0.1228070176	46.031145
	15	6	0.1224238007	22.695466		28	$(k-2)$	0.1228070176	48.009776

Table 4.1

k	s	λ_s	k	s	λ_s	k	s	λ_s
9	28 34	47.182765 59.051094	10	32 39	54.229641 68.087768	11	38 46	65.215856 81.088361
12	44 52	76.204964 92.095892	13	48 58	83.231405 103.100212	14	54 64	94.219572 114.102752
15	60 71	105.210117 127.097342	16	66 77	116.202593 138.099089	17	72 84	127.196419 151.094612
18	78 90	138.191614 162.096361	19	84 97	149.187721 175.092876	20	90 104	160.184757 188.090061

Table 4.2

or $j=k-4$ in which case ($j2$) is satisfied. The listed values were calculated to 16 significant figures by means of an electronic computer. However, once obtained it is relatively easy to check, if necessary by hand, that the corresponding ($j1$), ($j2$) or ($k-2$) are satisfied. Note that in each case when ($k-2$) is satisfied the optimising choice of θ is

$$\theta=\frac{2^{k-2}-1}{k2^{k-2}+1} \tag{4.5}$$

and then λ_s is given by

$$\lambda_s=(2s-2)\,\theta+1+\lambda_{s-1}(1-\theta). \tag{4.6}$$

The values given are all rounded up in the last decimal place.

We remark that for each k one further iteration will give

$$\lambda_{10}=15\ (k=5),\quad \lambda_{14}=22\ (k=6),\quad \lambda_{21}=35\ (k=7),\quad \lambda_{29}=50\ (k=8). \tag{4.7}$$

In Table 4.2 we extend Table 4.1 for selected values of s when $9\leqslant k\leqslant 20$. We make use of this in § 9.

We now treat the special cases $k=3$ and $k=4$.

THEOREM 4.3. *Let* $k=4$, $\lambda_1=1$, $\lambda_2=2$, $\lambda_3=13/4$.

$$\theta_4=2/(6+\sqrt{80}),\quad \lambda_4=6\theta_4+1+\lambda_3(1-\theta_4)=4.618033\ldots,$$

$$\theta_5=\frac{2}{11},\quad \lambda_5=8\theta_5+1+\lambda_4(1-\theta_5)=6.232936\ldots,$$

and suppose that $1 \leqslant s \leqslant 5$. *Then, provided that* $\lambda > \lambda_s$, $0 < \eta < \eta_0(\lambda - \lambda_s)$ *and* $P > P_0(\eta, s)$ *we have*

$$S_s(P, P^\eta) < P^\lambda.$$

Proof. When $s=1$ or 2 the result is classical. When $s=3$ or 4 the proof follows that of Theorem 4.1. Thus for $s=3$ we note that (j2) holds with $j=2$ and $\theta=1/8$ and for $s=4$ (j2) holds with $j=2$ and $\theta=\theta_4$.

When $s=5$ we begin by following the proof of Theorem 4.1. Thus (4.4) holds. Then we estimate $T_s(X, R, \theta)$ through the use of Lemma 3.6 with $\theta=\theta_5$. Hence

$$T_5(X, R, \theta) \ll X^{1+\theta} R^4 S_4(X^{1-\theta}, R) + X^{6/5-(11/5)\theta+1/15+\varepsilon^2} R S_5(X^{1-\theta}, R)^{4/5}.$$

Therefore we may conclude the argument by following the proof of Theorem 4.1 in the case $(k-2)$. We need only check that λ_5 satisfies

$$\lambda_5 \geqslant 6 + \frac{11}{57} = 6 + \frac{1}{15} \Big/ \left(\frac{1}{5} + \frac{2}{11} - \frac{2}{55} \right).$$

THEOREM 4.4. *Let* $k=3$ *and suppose that* $\lambda > 13/4$. *Then, provided that* $0 < \eta < \eta_0(\lambda - 13/4)$ *and* $P > P_0(\eta)$ *we have*

$$S_3(P, P^\eta) < P^\lambda.$$

Proof. We again follow the proof of Theorem 4.1, so that (4.4) holds, but we take $\theta = 1/8$ and estimate $T_3(X, R, \theta)$ through the use of Lemma 3.7. Thus

$$X^{3\theta} T_3(X, R, \theta) \ll X^{3+2\theta+\varepsilon(1-\theta)} R + X^{\frac{7}{6}+\frac{3}{2}\theta+\varepsilon^2} R S_3(X^{1-\theta}, R)^{\frac{2}{3}}.$$

Again the proof is completed as before. We need only note that

$$3 + 2\theta = \frac{13}{4} \quad \text{and} \quad \frac{7}{6} + \frac{3}{2}\theta + \frac{2}{3}\left(\frac{13}{4}(1-\theta) \right) = \frac{13}{4}.$$

5. The estimation of $G(k)$ when $4 \leqslant k \leqslant 8$

We establish Theorems 1.1. and 1.2 through the medium of the Hardy-Littlewood method. We consider the representation of a large natural number n in the form

$$x_1^k + \ldots + x_{2t}^k + y_1^k + \ldots + y_u^k = n$$

where each x_i has the property that it has no prime factor exceeding n^η where η is a sufficiently small but fixed positive number. This enables us to combine Theorem 4.2 with Weyl's inequality on the minor arcs.

The restriction on the x_i is unlike any condition that has been used hitherto in connection with Waring's problem. Moreover u is usually too small for a direct application of classical methods on the major arcs. Thus it is necessary to develop a new technique for dealing with the major arcs.

We suppose now that n is large, and P and W satisfy

$$P = n^{1/k}, \quad 2 \leq W \leq P. \tag{5.1}$$

Let

$$\mathfrak{M}(q, a) = \{\alpha\colon |\alpha - a/q| \leq (2kq)^{-1}P^{1-k}\} \tag{5.2}$$

denote a typical major arc, let $\mathfrak{M}$ denote the union of the $\mathfrak{M}(q, a)$ with $1 \leq a \leq q \leq P$, $(a, q) = 1$, and let

$$\mathfrak{m} = ((2k)^{-1}P^{1-k}, 1 + (2k)^{-1}P^{1-k}] \setminus \mathfrak{M} \tag{5.3}$$

denote the corresponding minor arcs. Clearly the $\mathfrak{M}(q, a)$ are disjoint and contained in $((2k)^{-1}P^{1-k}, 1 + (2k)^{-1}P^{1-k}]$.

It is not possible to estimate precisely the bulk of our generating functions throughout $\mathfrak{M}$ without developing considerable machinery to handle the distribution of the elements of $\mathscr{A}(P, P^\eta)$ in arithmetic progressions to relatively large moduli, at least in mean. We therefore adopt a procedure for pruning the major arcs.

Let

$$\mathfrak{N}(q, a) = \{\alpha\colon |\alpha - a/q| \leq (2kq)^{-1}WP^{-k}\}, \tag{5.4}$$

and let $\mathfrak{N}$ denote the union of the $\mathfrak{N}(q, a)$ with $1 \leq a \leq q \leq W$, $(a, q) = 1$. Evidently $\mathfrak{N}(q, a) \subset \mathfrak{M}(q, a)$ and $\mathfrak{N} \subset \mathfrak{M}$.

Let

$$\mathfrak{n} = ((2k)^{-1}P^{1-k}, 1 + (2k)^{-1}P^{1-k}] \setminus \mathfrak{N} \tag{5.5}$$

so that

$$\mathfrak{n} = (\mathfrak{M} \setminus \mathfrak{N}) \cup \mathfrak{m}.$$

We first record some useful information regarding the standard generating function

$$f(\alpha)=\sum_{x\leqslant P} e(\alpha x^k). \tag{5.6}$$

By Theorem 4.1 of [Va2], for $\alpha\in\mathfrak{M}(q,a)$ we have

$$f(\alpha)=V(\alpha,q,a)+O(q^{\frac{1}{2}+\varepsilon}) \tag{5.7}$$

where

$$V(\alpha,q,a)=q^{-1}S(q,a)\,v(\alpha-a/q), \tag{5.8}$$

$$S(q,a)=\sum_{r=1}^{q} e(ar^k/q) \tag{5.9}$$

and

$$v(\beta)=\sum_{x\leqslant n}\frac{1}{k}x^{\frac{1}{k}-1}e(\beta x). \tag{5.10}$$

Moreover, by Lemma 2.8 of [Va2] we have

$$v(\beta)\ll\min(P,\|\beta\|^{-1/k}). \tag{5.11}$$

We will find the following lemma particularly useful.

LEMMA 5.1. *Suppose that $k\geqslant 3$ and $s\geqslant k+2$. Then*

$$\int_{\mathfrak{M}}|f(\alpha)|^s d\alpha\ll P^{s-k}$$

and

$$\int_{\mathfrak{M}\setminus\mathfrak{N}}|f(\alpha)|^s d\alpha\ll W^{\varepsilon-1/k}P^{s-k}.$$

Proof. By (5.7), for $\alpha\in\mathfrak{M}(q,a)$ we have

$$|f(\alpha)|^s-|V(\alpha,q,a)|^s\ll\left(q^{\frac{1}{2}+\varepsilon}\right)^s+q^{\frac{1}{2}+\varepsilon}|V(\alpha,q,a)|^{s-1}.$$

Let

$$V(\alpha)=V(\alpha,q,a)\quad(\alpha\in\mathfrak{M}(q,a),\ 1\leqslant a\leqslant q\leqslant P,\ (a,q)=1)$$

and $\mathcal{M}=\mathfrak{M}$ or $\mathfrak{M}\setminus\mathfrak{N}$. Then the argument of Theorem 4.4 of [Va2] establishes that

$$\int_{\mathcal{M}} |f(\alpha)|^s d\alpha = \int_{\mathcal{M}} |V(\alpha)|^s d\alpha + O(P^{s-\frac{1}{2}-k+\varepsilon}).$$

Moreover, by (5.11) and in the notation of Lemma 4.9 of [Va2] we have

$$\int_{\mathcal{M}} |V(\alpha)|^s d\alpha \ll P^{s-k} \sum_{q\leqslant P} S_s^*(q) \min\left(\left(\frac{q}{Y}\right)^{\frac{s}{k}-1}, 1\right)$$

where $Y=1$ when $\mathcal{M}=\mathfrak{M}$ and $Y=W$ when $\mathcal{M}=\mathfrak{M}\setminus\mathfrak{N}$. By a variant of the argument of Lemma 4.9 of [Va2] we have, for $s\geqslant k+2$,

$$\sum_{q\leqslant Z} q^{(s-1-k)/k} S_s^*(q) \ll Z^{\varepsilon}.$$

Therefore

$$\sum_{q\leqslant Y} q^{(s-k)/k} S_s^*(q) \ll Y^{\varepsilon+1/k}$$

and

$$\sum_{q\geqslant Y} S_s^*(q) \ll Y^{\varepsilon-(s-1-k)/k}.$$

The lemma now follows easily.

It is also useful to record the standard estimate for f on $\mathfrak{m}$ that is a consequence of Weyl's inequality (Lemma 2.4 of [Va2]), namely that

$$f(\alpha) \ll P^{1-\sigma+\varepsilon} \quad (\alpha \in \mathfrak{m}) \tag{5.12}$$

where

$$\sigma = 2^{1-k}.$$

Henceforward we suppose that η is a sufficiently small but fixed positive number, that $n>n_0(\eta)$, and take

$$R = P^{\eta} \tag{5.13}$$

k	$t(k)$	$u(k)$	$v(k)$
4	5	1	2
5	9	1	1
6	13	2	3
7	20	1	1
8	28	1	2

Table 5.1

and

$$g(\alpha) = \sum_{x \in \mathcal{A}(P,R)} e(\alpha x^k). \tag{5.14}$$

When $4 \leqslant k \leqslant 8$ we suppose that $t=t(k)$, $u=u(k)$ and $v=v(k)$ are given by Table 5.1.

LEMMA 5.2. *Let t, u, v be as in Table 5.1, and let*

$$L = \int_0^1 |f(\alpha)^{2u} g(\alpha)^{2t}| \, d\alpha$$

and

$$I(\mathfrak{B}) = \int_{\mathfrak{B}} |f(\alpha)^{v} g(\alpha)^{2t}| \, d\alpha.$$

Then

$$L \ll P^{2t+2u-k},$$

$$I([0,1]) \ll P^{2t+v-k}$$

and there is a positive number δ such that

$$I(\mathfrak{n}) \ll P^{2t+v-k} W^{-\delta}.$$

Proof. By Table 4.1,

$$\lambda_t + v(1-2^{1-k}) < 2t+v-k.$$

Therefore, by Theorems 4.2 and 4.3 and (5.12) we have

$$\int_{\mathfrak{m}} |f(\alpha)^{2u} g(\alpha)^{2t}| \, d\alpha \ll P^{2t+2u-k-\delta}, \quad I(\mathfrak{m}) \ll P^{2t+v-k-\delta}. \tag{5.15}$$

Note that, by Table 5.1 we have $v \leqslant 2u$.

By Hölder's inequality

$$\int_{\mathfrak{M}} |f(\alpha)^{2u} g(\alpha)^{2t}|\, d\alpha \leqslant \left(\int_{\mathfrak{M}} |f(\alpha)|^{2t+2u} d\alpha\right)^{\frac{u}{t+u}} \left(\int_{\mathfrak{M}} |g(\alpha)|^{2t+2u} d\alpha\right)^{\frac{t}{t+u}}.$$

By replacing the last integral over $\mathfrak{M}$ by one over $[0, 1]$ and interpreting the result in terms of the underlying diophantine equation we see that

$$\int_{\mathfrak{M}} |g(\alpha)|^{2t+2u} d\alpha \leqslant L.$$

Moreover, by Table 5.1 we have $2t+2u \geqslant k+2$. Therefore, by (5.15) and Lemma 5.1 we obtain

$$L \ll P^{2t+2u-k} + (P^{2t+2u-k})^{\frac{u}{t+u}} L^{\frac{t}{t+u}}$$

and so secure the first part of the lemma.

Let $\mathcal{M} = \mathfrak{M}$ or $\mathfrak{M} \setminus \mathfrak{N}$. Then, by Hölder's inequality once more, we have

$$I(\mathcal{M}) \leqslant \left(\int_{\mathcal{M}} |f(\alpha)|^{\frac{v}{u}(t+u)} d\alpha\right)^{\frac{u}{t+u}} \left(\int_{\mathcal{M}} |g(\alpha)|^{2t+2u} d\alpha\right)^{\frac{t}{t+u}}.$$

As before the last integral here is bounded by L. By Table 5.1, $v \geqslant u$ and $t+u \geqslant k+2$. Hence, by Lemma 5.1 and (5.15),

$$I([0, 1]) \ll P^{2t+v-k} + \left(P^{\frac{v}{u}(t+u)-k}\right)^{\frac{u}{t+u}} (P^{2t+2u-k})^{\frac{t}{t+u}}$$
$$\ll P^{2t+v-k}$$

and

$$I(\mathfrak{n}) \ll P^{2t+v-k-\delta} + \left(P^{\frac{v}{u}(t+u)-k} W^{\varepsilon-1/k}\right)^{\frac{u}{t+u}} (P^{2t+2u-k})^{\frac{t}{t+u}}$$
$$\ll P^{2t+v-k} W^{-\delta}$$

as required.

The next step in our argument is to estimate

$$\int_{\mathfrak{N}} f(\alpha)^v g(\alpha)^{2t} e(-\alpha n)\, d\alpha$$

asymptotically for a suitable choice of W. We achieve this through the approximation for g embodied in Lemma 5.4. Before starting this we introduce some notation. Let $\varrho(x)$ denote Dickman's function, defined for real x by

$$\varrho(x) = 0 \text{ when } x<0,$$
$$\varrho(x) = 1 \text{ when } 0\leqslant x\leqslant 1,$$
$$\varrho \text{ continuous for } x>0,$$
$$\varrho \text{ differentiable for } x>1,$$
$$x\varrho'(x) = -\varrho(x-1) \text{ when } x>1.$$

For an extensive study of the properties of ϱ see [B1]. Note that for $x\geqslant 0$, $\varrho(x)$ is positive and decreasing.

We further define

$$w(\beta) = \sum_{R^k<m\leqslant n} \frac{1}{k} m^{\frac{1}{k}-1} \varrho\left(\frac{\log m}{k\log R}\right) e(\beta m) \tag{5.16}$$

and

$$W(\alpha, q, a) = q^{-1}S(q, a)\, w(\alpha-a/q). \tag{5.17}$$

At several stages in our arguments we require some knowledge of the asymptotics of $\mathscr{A}(P, R)$. This is summarised in

LEMMA 5.3. *Let τ be a fixed positive number and suppose that $R\leqslant X\leqslant R^{\tau}$. Then*

$$\operatorname{card}(\mathscr{A}(X, R)) = X\varrho\left(\frac{\log X}{\log R}\right)+O\left(\frac{X}{\log X}\right).$$

Proof. The lemma is immediate from (1.3) and (1.4) of [B2] and standard estimates from prime number theory.

LEMMA 5.4. *Suppose that $q\leqslant R$, $(a, q)=1$ and $\beta=\alpha-a/q$. Then*

$$g(\alpha) = W(\alpha, q, a)+O\left(\frac{qP}{\log P}(1+n|\beta|)\right)$$

and

$$W(\alpha, q, a) \ll q^{-1/k}\min(P, \|\beta\|^{-1/k}).$$

Proof. Let Π denote the product of all primes p with $R<p\leqslant P$, and suppose that $R<m\leqslant P$. Then

$$\sum_{\substack{x\in\mathscr{A}(m,R)\\ x\equiv r\ (\mathrm{mod}\ q)}} 1=\sum_{\substack{d\mid\Pi\\ d\leqslant m}}\mu(d)\sum_{\substack{y\leqslant m/d\\ yd\equiv r\ (\mathrm{mod}\ q)}} 1.$$

Since $d\mid\Pi$ and $q\leqslant R$ we have $(d,q)=1$. Hence

$$\sum_{\substack{x\in\mathscr{A}(m,R)\\ x\equiv r\ (\mathrm{mod}\ q)}} 1=\sum_{\substack{d\mid\Pi\\ d\leqslant m}}\mu(d)\left(\frac{m}{dq}+O(1)\right)$$
$$=\frac{1}{q}\sum_{\substack{d\mid\Pi\\ d\leqslant m}}\mu(d)\frac{m}{d}+O\Bigg(\sum_{\substack{d\mid\Pi\\ d\leqslant m}}1\Bigg).$$

Hence, by the case $q=1$, we obtain

$$\sum_{\substack{x\in\mathscr{A}(m,R)\\ x\equiv r\ (\mathrm{mod}\ q)}} 1=\frac{1}{q}\sum_{x\in\mathscr{A}(m,R)}1+O\Bigg(\sum_{\substack{d\mid\Pi\\ d\leqslant m}}1\Bigg).$$

The error term here is bounded by the number of natural numbers not exceeding m which are coprime with $\Pi_{p\leqslant R}\,p$. Hence, by Theorem 2.2 of [HR] and elementary prime number theory we have

$$\sum_{\substack{x\in\mathscr{A}(m,R)\\ x\equiv r\ (\mathrm{mod}\ q)}} 1=\frac{1}{q}\sum_{x\in\mathscr{A}(m,R)}1+O\left(\frac{P}{\log P}\right)\quad(R<m\leqslant P).$$

Therefore

$$\sum_{x\in\mathscr{A}(m,R)} e(ax^k/q)=q^{-1}S(q,a)\sum_{x\in\mathscr{A}(m,R)}1+O\left(\frac{qP}{\log P}\right).$$

Let

$$S_y=\sum_{x\in\mathscr{A}(y^{1/k},R)}(e(ax^k/q)-q^{-1}S(q,a))$$

so that for $R^k<y\leqslant n$ we have

$$S_y\ll\frac{qP}{\log P}.$$

Then, by partial summation

$$\sum_{\substack{x\in\mathscr{A}(P,R)\\ x>R}} (e(\alpha x^k)-q^{-1}S(q,a)\,e(\beta x^k))$$

$$=\sum_{R^k<y\leq n} S_y(e(\beta y)-e(\beta(y+1)))+S_n e(\beta(n+1))-S_{[R^k]}\,e(\beta([R^k]+1)).$$

Hence

$$\sum_{\substack{x\in\mathscr{A}(P,R)\\ x>R}} (e(\alpha x^k)-q^{-1}S(q,a)\,e(\beta x^k))\ll\frac{nqP}{\log P}|\beta|+\frac{qP}{\log P}. \tag{5.18}$$

Let

$$T_y=\sum_{x\in\mathscr{A}(y^{1/k},R)} 1.$$

Then, by the previous lemma

$$T_y=y^{1/k}\varrho\left(\frac{\log y}{\log R}\right)+O\left(\frac{y^{1/k}}{\log P}\right)$$

whenever $R^k<y\leq n$. Hence, by partial summation

$$\sum_{x\in\mathscr{A}(P,R),\,x>R} e(\beta x^k)=\sum_{R^k<y\leq n} T_y(e(\beta y)-e(\beta(y+1)))+T_n\,e(\beta(n+1))-T_{[R^k]}\,e(\beta([R^k]+1))$$

$$=\sum_{R^k<y\leq n}\left(y^{1/k}\varrho\left(\frac{\log y}{k\log R}\right)-(y-1)^{1/k}\varrho\left(\frac{\log(y-1)}{k\log R}\right)\right)e(\beta y)$$
$$+O\left(\frac{P}{\log P}(1+n|\beta|)\right).$$

When $y>R^k+1$, an application of the mean value theorem shows, since ϱ' is bounded, that

$$y^{1/k}\varrho\left(\frac{\log y}{k\log R}\right)-(y-1)^{1/k}\varrho\left(\frac{\log(y-1)}{k\log R}\right)=\frac{1}{k}y^{\frac{1}{k}-1}\varrho\left(\frac{\log y}{k\log R}\right)+O\left(\frac{y^{\frac{1}{k}-1}}{\log R}+\frac{1}{k}y^{\frac{1}{k}-2}\right).$$

Thus, by (5.16),

$$\sum_{\substack{x\in\mathcal{A}(P,R)\\ x>R}} e(\beta x^k) = w(\beta)+O\left(\frac{P}{\log P}(1+n|\beta|)\right).$$

This with (5.18) establishes the first part of the lemma.

The second part of the lemma follows by the methods of Lemmas 2.8 and 4.6 of [Va2] and the monotonicity of ϱ.

By imitating the usual method of estimation on the major arcs, where necessary using Lemma 5.4, we obtain when $W\leqslant R$

$$\int_{\mathfrak{N}} f(\alpha)^v g(\alpha)^{2t} e(-\alpha n)\,d\alpha = \mathfrak{S}(n)\,J(n)+O\left(P^{2t+v-k}\left(\frac{W^C}{\log P}+W^{-\delta}\right)\right) \tag{5.19}$$

where C and δ are positive constants that depend only on t, v and k, where $\mathfrak{S}(n)$ is the usual singular series in Waring's problem

$$\mathfrak{S}(n) = \sum_{q=1}^{\infty} \sum_{\substack{a=1\\ (a,q)=1}}^{q} (S(q,a)/q)^{2t+v} e(-an/q),$$

and where

$$J(n) = \sum_{y_1}\dots\sum_{y_v}\sum_{x_1}\dots\sum_{x_{2t}} k^{-v-2t}(y_1\dots y_v x_1\dots x_{2t})^{\frac{1}{k}-1}\varrho_1\dots\varrho_{2t}$$

with

$$\varrho_j = \varrho\left(\frac{\log x_j}{k\log R}\right)$$

and the multiple sum is over $y_1,\dots,y_v$, $x_1,\dots,x_{2t}$ satisfying

$$y_1\leqslant n,\ \dots,\ y_v\leqslant n,\quad R^k<x_1\leqslant n,\ \dots,\ R^k<x_{2t}\leqslant n,$$

$$y_1+\dots+y_v+x_1+\dots+x_{2t} = n.$$

By Lemma 5.2 and (5.19) with W a suitable power of $\log P$ we obtain

$$R(n) = \mathfrak{S}(n)\,J(n)+O(P^{2t+v-k}(\log P)^{-v})$$

where v is a positive constant and $R(n)$ is the number of solutions of

$$y_1^k+\ldots+y_v^k+x_1^k+\ldots+x_{2t}^k = n$$

with $x_j \in \mathcal{A}(P,R)$.

A simple counting argument combined with the fact that

$$\varrho\left(\frac{\log x_j}{k \log R}\right) \gg 1$$

when $R^k < x_j \leq n$ establishes that

$$J(n) \gg P^{2t+v-k}.$$

Moreover, by Theorem 4.6 of [Va2] and Table 5.1,

$$\mathfrak{S}(n) \gg 1$$

when $k \geq 5$. This conclusion is also evident when $k=4$ and $n \equiv r \pmod{16}$ with $1 \leq r \leq 12$ by the argument in the penultimate paragraph on page 87 of [Va2].

This establishes Theorems 1.1 and 1.2.

6. Sums of three kth powers

Theorem 1.3 follows immediately from Theorem 4.4 *via* Cauchy's inequality and the lower bound

$$\text{card}(\mathcal{A}(P,P^\eta)) \gg P$$

that is a consequence of Lemma 5.4, for example.

Theorem 1.4 follows likewise from Theorems 4.2 and 4.3 on observing that in Theorem 4.2 if $\lambda_3 = 3+1/k$, then $(j1)$ holds with $j=2$, $\theta = 1/2k$.

7. A simplified estimation and an exponential sum

We obtain Theorem 1.5 by combining our method with an idea of Vinogradov. We first state a bound for $S_s(P,P^\eta)$ that avoids an excess of calculation when k is large.

THEOREM 7.1. *Suppose that* $k \geq 5$, $\lambda_1 = 1$ *and that for* $s \geq 2$, λ_s *is given by*

$$\lambda_s = 2s-k+(k-2)\left(1-\frac{1}{k}\right)^{s-2}.$$

Suppose further that $\lambda > \lambda_s$. *Then, provided that* $0 < \eta < \eta_0(\lambda - \lambda_s)$ *and* $P > P_0(\eta, s)$, *we have*

$$S_s(P, P^\eta) < P^\lambda.$$

Proof. The theorem follows from Theorem 4.2 on observing that (j1) is satisfied with $\theta = 1/k$ when $t \leq 2^{k-1}$ and that ($k-2$) is also satisfied with $\theta = 1/k$ when $t \geq 2k-2$.

Let

$$X = P^{1/2} \tag{7.1}$$

and for a sufficiently small positive number η let

$$\mathscr{B} = \left\{ x : x = py, \frac{1}{2}X < p \leq X, y \in \mathscr{A}(X, X^\eta) \right\} \tag{7.2}$$

and

$$h(\alpha) = \sum_{x \in \mathscr{B}} e(\alpha x^k). \tag{7.3}$$

By Hölder's inequality,

$$|h(\alpha)|^{2s} \leq X^{2s-1} \sum_{\frac{1}{2}X < p \leq X} \left| \sum_{y \leq Y} b_y e(\alpha p^k y) \right|^2 \tag{7.4}$$

where $Y = sX^k$ and b_y is the number of solutions of

$$y_1^k + \ldots + y_s^k = y$$

with $y_j \in \mathscr{A}(X, X^\eta)$. We note for future reference that

$$\sum_y b_y^2 = S_s(X, X^\eta). \tag{7.5}$$

Let $\mathfrak{m}$ be as in the hypothesis of Theorem 1.5, and let $\alpha \in \mathfrak{m}$. Choose a, q so that $(a, q) = 1$, $q \leq 2X^k$, $|\alpha - a/q| \leq \frac{1}{2} q^{-1} X^{-k}$. Then, by the definition of $\mathfrak{m}$, either $q > X$, or $q \leq X$ and $|\alpha - a/q| > q^{-1} X^{1-2k} \gg q^{-1} X^{1-k} Y^{-1}$. Thus the hypothesis of Lemma 5.4 of [Va2] is satisfied. Hence we may estimate the right hand side of (7.4) in the same way that (5.44) of [Va2] is estimated. Thus

$$|h(\alpha)|^{2s} \ll X^{2s-1} Y^{1+\varepsilon} \sum_y |b_y|^2.$$

Therefore, by (7.5) and Theorem 7.1, when $0<\eta<\eta_0(\varepsilon)$ and $P>P_0(\eta, s)$, we have

$$|h(\alpha)|^{2s} \ll X^{2s-1+k+\lambda_s+\varepsilon}.$$

This means that

$$h(\alpha) \ll P^{1-\sigma+\varepsilon}$$

where

$$\sigma = \sigma(k, s) = \frac{1}{4s}\left(1-(k-2)\left(1-\frac{1}{k}\right)^{s-2}\right).$$

In addition, by (7.1), (7.2) and Lemma 5.3 we have

$$\operatorname{card} \mathscr{B} \gg \frac{X^2}{\log X} \gg \frac{P}{\log P}.$$

This establishes the main part of Theorem 1.5.

The maximum of $\sigma(k, s)$ as s varies is attained for a value of s satisfying

$$\left|s-\lambda\left(\log\frac{k}{k-1}\right)^{-1}\right| < 1$$

where λ is the larger root of the transcendental equation

$$(\lambda+1)\frac{k^2(k-2)}{(k-1)^2} = e^{\lambda}. \tag{7.6}$$

Thus

$$\lambda = \log k+\log\log k+O\left(\frac{\log\log k}{\log k}\right) \tag{7.7}$$

and

$$\varrho(k) = \frac{\log\frac{k}{k-1}}{4(\lambda+1)}\left(1+O\left(\frac{1}{k\log k}\right)\right). \tag{7.8}$$

This completes the proof of Theorem 1.5.

8. The estimation of $G(k)$ when k is large

We now investigate the possibility of combining Theorems 1.5 and 7.1 through a variant of the argument developed in § 5 to establish Theorem 1.1.

Let $\mathfrak{M}$, $\mathfrak{N}$, $\mathfrak{m}$, $\mathfrak{n}$ be as in § 5, and let f, g, h be as in (5.7), (5.14) and (7.3). We suppose that $0<\eta<\eta_0(\varepsilon)$, that

$$u \geqslant k+1, \tag{8.1}$$

$$2t\varrho(k) > (k-2)\left(1-\frac{1}{k}\right)^{u-2} \tag{8.2}$$

where $\varrho(k)$ is as in Theorem 1.5 and define

$$L = \int_0^1 |f(\alpha)^2 g(\alpha)^{2u} h(\alpha)^{2t}|\, d\alpha$$

and

$$I = \int_{\mathfrak{n}} |f(\alpha)\, g(\alpha)^{2u} h(\alpha)^{2t}|\, d\alpha.$$

By Hölder's inequality and Theorem 1.5 (note that $\mathfrak{m} \subset m$),

$$L \ll \int_{\mathfrak{m}} P^{2+2t-2t\varrho(k)+\varepsilon} |g(\alpha)|^{2u} d\alpha + \left(\int_{\mathfrak{M}} |f(\alpha)|^{k+2} d\alpha\right)^{\frac{2}{k+2}} \left(\int_0^1 |g(\alpha)^{2u} h(\alpha)^{2t}|^{\frac{k+2}{k}}\right)^{\frac{k}{k+2}}.$$

By Theorem 7.1 and (8.2) the first integral on the right is

$$\ll P^{2+2t+2u-k}.$$

By (8.1), $4u/k \geqslant 2$ so that the last integral on the right is bounded by

$$\int_0^1 |g(\alpha)^{2+2u} h(\alpha)^{2t}|\, P^{\frac{4u}{k}-2+\frac{4t}{k}} d\alpha \leqslant L P^{(4u+4t-2k)/k}.$$

Hence, by Lemma 5.1,

$$L \ll P^{2+2t+2u-k} + (P^2)^{\frac{2}{k+2}} (L P^{(4u+4t-2k)/k})^{\frac{k}{k+2}},$$

whence

$$L \ll P^{2+2t+2u-k}.$$

Now a cognate argument gives

$$I \ll P^{1+2t+2u-k-\delta} + \left(\int_{\mathfrak{M} \setminus \mathfrak{N}} |f(\alpha)|^{k+2} d\alpha\right)^{\frac{1}{k+2}} \left(\int_0^1 |g(\alpha)^{2u} h(\alpha)^{2t}|^{\frac{k+2}{k+1}} d\alpha\right)^{\frac{k+1}{k+2}}$$

where δ is a suitable postive number. Since $u \geqslant k+1$ the last integral on the right is bounded by

$$LP^{\frac{2u}{k+1}-2+\frac{2t}{k+1}}.$$

By applying our bound for L and appealing to Lemma 5.1 we obtain

$$I \ll P^{1+2t+2u-k}W^{-\delta}.$$

The process developed for dealing with the contribution from $\mathfrak{N}$ in § 5 now shows that every large n can be written as the sum of $1+2t+2u$ kth powers. We take

$$v=u-2, \quad t=1+\left[\frac{k-2}{2\varrho(k)}\left(1-\frac{1}{k}\right)^{v}\right].$$

Thus

$$G(k) \leqslant 7+2v+2\left[\frac{k-2}{2\varrho(k)}\left(1-\frac{1}{k}\right)^{v}\right].$$

The remark after Theorem 1.6 can be justified by observing that the optimising choice of v occurs with $|v-\mu|<1$ where

$$\mu \log\frac{k}{k-1} = \log\left(\frac{k-2}{2\varrho(k)}\log\frac{k}{k-1}\right).$$

Thus

$$G(k) \leqslant \frac{2\log\left(e\frac{k-2}{2\varrho(k)}\log\frac{k}{k-1}\right)}{\log\frac{k}{k-1}} + O(\log k)$$
$$= 2k\log\left(\frac{e}{2\varrho(k)}\right) + O(\log k),$$

and by (7.6), (7.7) and (7.8)

$$\log\frac{1}{4\varrho(k)} = \log(k(\lambda+1)) + O\left(\frac{1}{k}\right)$$
$$= \log(k\log k) + O\left(\frac{\log\log k}{\log k}\right).$$

k	$s(k)$	k	$s(k)$	k	$s(k)$	k	$s(k)$
8	22	12	44	15	60	18	78
9	28	13	48	16	66	19	84
10	32	14	54	17	72	20	90
11	38						

Table 9.1

9. The estimation of $G(k)$ when $9\leqslant k\leqslant 20$

When k is of moderate size we may vary the argument of the previous section by using Theorem 4.2 with optimal choices for the λ_s to establish improved versions of both Theorem 1.6 and Theorem 1.7. A further improvement can be brought about by employing a more precise form of the proof of Theorem 1.6.

Let

$$X = P^{\frac{k}{2k-1}}, \tag{9.1}$$

$$Z = PX^{-1} \tag{9.2}$$

and define the generating function h by

$$\mathscr{C} = \left\{x\colon x = pz,\ \frac{1}{2}X<p\leqslant X,\ z\in\mathscr{A}(Z,Z^{\eta})\right\}, \tag{9.3}$$

$$h(\alpha) = \sum_{x\in\mathscr{C}} e(\alpha x^k) \tag{9.4}$$

We now define $s=s(k)$ as in Table 9.1.

Then, by Lemma 4.2 and Tables 4.1 and 4.2 we have

$$S_s(Z,Z^{\eta}) < Z^{\lambda} \tag{9.5}$$

whenever $\lambda>\lambda_s$, $0<\eta<\eta_0(\lambda-\lambda_s)$ and $P>P_0(\eta)$.

Since s is even we may write $2r=s$. By Hölder's inequality

$$\begin{aligned} |h(\alpha)|^s &\leqslant X^{s-1} \sum_{\frac{1}{2}X<p\leqslant X} \left| \sum_{z\in\mathscr{A}(Z,Z^{\eta})} e(\alpha p^k z^k)\right|^{2r} \\ &= X^{s-1} \sum_{\frac{1}{2}X<p\leqslant X} \sum_{|y|\leqslant Y} c_y e(\alpha p^k y) \end{aligned} \tag{9.6}$$

k	$\sigma(k)$	k	$\sigma(k)$	k	$\sigma(k)$	k	$\sigma(k)$
8	0.01008306	12	0.00481491	15	0.00346504	18	0.00270010
9	0.00791794	13	0.00425964	16	0.00316742	19	0.00251303
10	0.00652403	14	0.00382220	17	0.00291609	20	0.00234894
11	0.00553974						

Table 9.2

where c_y is the number of solutions of

$$z_1^k - z_2^k + \ldots + z_{2r-1}^k - z_{2r}^k = y$$

with $z_i \in \mathscr{A}(Z, Z^\eta)$ and

$$Y = rZ^k. \tag{9.7}$$

Let $\mathfrak{n}$ denote the set of real numbers α such that whenever $a \in \mathbf{Z}$, $q \in \mathbf{N}$, $(a, q)=1$ and $|\alpha - a/q| \leqslant q^{-1}X^{1-k}Y^{-1}$ one has $q > X$.

Let $\alpha \in \mathfrak{n}$ and choose a, q so that $(a, q)=1$, $q \leqslant 2(2X)^k$ and $|\alpha - a/q| \leqslant q^{-1}2^{-1}(2X)^{-k}$. Then, by the definition of $\mathfrak{n}$, either $q > X$, or $q \leqslant X$ and $|\alpha - a/q| > q^{-1}X^{1-k}Y^{-1}$. We now appeal to a variant of Vinogradov's estimate for sums of the kind on the right of (9.6). The most effective form for the purpose at hand is that contained in the main Lemma of [T2]. Thus, as $c_y = c_{-y}$,

$$\sum_{\frac{1}{2}X<p\leqslant X} \sum_{|y|\leqslant Y} c_y e(\alpha p^k y) = \sum_{\frac{1}{2}X<p\leqslant X} \left(c_0 + 2\,\mathrm{Re} \sum_{0<y\leqslant Y} c_y e(\alpha p^k y) \right)$$

$$\ll X^\varepsilon (XY + X^k) \left(\sum_{|y|\leqslant Y} c_y^2 \right)^{1/2}.$$

Hence, by (9.1), (9.2), (9.6) and (9.7)

$$|h(\alpha)|^{2s} \ll X^{2s+k-2+\varepsilon} S_s(Z, Z^\eta).$$

Therefore, by (9.5),

$$h(\alpha) \ll P^{1-\sigma+\varepsilon} \quad (0<\eta<\eta_1(\varepsilon),\ P>P_1(\eta))$$

k	$u(k)$	$t(k)$	k	$u(k)$	$t(k)$	k	$u(k)$	$t(k)$	k	$u(k)$	$t(k)$
9	34	6	12	52	20	15	71	28	18	90	36
10	39	14	13	58	24	16	77	32	19	97	37
11	46	16	14	64	27	17	84	33	20	104	39

Table 9.3

where

$$\sigma=\sigma(k)=\frac{(k-1)(2s-\lambda_s)-k(k-2)}{2s(2k-1)}.$$

The values of σ that arise from Tables 4.1 and 4.2 by taking s as in Table 9.1 are listed below in Table 9.2. The values given are rounded down in the last decimal place.

Now let $u=u(k)$ and $t=t(k)$ be given by Table 9.3.

We take $\mathfrak{M}$, $\mathfrak{N}$, $\mathfrak{m}$, $\mathfrak{n}$ as in § 5, so that $\mathfrak{m}\subset\mathfrak{n}$ and define $t_0=t_0(k)$ by

$$t_0-1<\frac{1}{2}t\leqslant t_0.$$

By Tables 4.2, 9.2 and 9.3

$$\lambda_u+t(1-\sigma)<2u+t-k \quad (k\neq 9 \text{ or } 15).$$

Therefore, by a variant of the argument of § 8, and with f and g as in (5.7) and (5.14) we have

$$\int_0^1 |f(\alpha)^2 g(\alpha)^{2u} h(\alpha)^{2t_0}|\,d\alpha \ll P^{2+2u+2t_0-k}$$

and

$$\int_{\mathfrak{n}} |f(\alpha)\,g(\alpha)^{2u} h(\alpha)^t|\,d\alpha \ll P^{1+2u+t-k}W^{-\delta}. \tag{9.8}$$

When $k=9$ or 15 we observe that by Lemma 2.4 and Theorem 5.3 of [Va2]

$$f(\alpha)\ll P^{1-\tau} \quad (\alpha\in\mathfrak{m})$$

where

$$\tau = 0.00390625 \ (k=9), \quad \tau = 0.000347551 \ (k=15).$$

Thus

$$1-\tau+\lambda_u+t(1-\sigma) < 1+2u+t-k$$

and our argument may proceed as before. Thus (9.8) holds in this case also.

It now follows by a kindred method to that used in § 5 for dealing with $\mathfrak{N}$ that

$$G(k) \leqslant 1+2u+t$$

and this establishes Theorem 1.7.

10. Another exponential sum

We now proceed with the proof of Theorem 1.8. Let $\alpha \in \mathfrak{m}$ and choose a, q so that

$$(a,q)=1, \quad q \leqslant P^{k/2}, \quad |\alpha - a/q| \leqslant q^{-1}P^{-k/2}. \tag{10.1}$$

We desire to convert $S(\alpha)$ into the kind of sum considered in the proof of Theorem 1.5. However the possibility that (q,x^k) may be large is a nuisance. We deal with this by first removing the common factors that may arise by writing

$$S(\alpha) = \sum_{q_0|q} \sum_{\substack{x\in\mathcal{A}(P,R)\\ (q,x^k)=q_0}} e(\alpha x^k).$$

Let q_k^k be the largest kth power dividing q_0, q_{q-1}^{k-1} be the largest $(k-1)$st power dividing $q_0 q_k^{-k}$, and so on. Then $q_1 \dots q_k | x$. Hence

$$\begin{aligned} S(\alpha) &= \sum_{q_0|q} \sum_{\substack{yq_1\dots q_k\in\mathcal{A}(P,R)\\ (q/q_0,\, y^k q_1^{k-1}\dots q_{k-1})=1}} e(\alpha y^k q_1^k \dots q_k^k). \\ &= {\sum_{q_0 r=q}}^{*} T(\alpha q_1^k \dots q_k^k, P/(q_1\dots q_k), R, r) + O(q^\varepsilon P^{1-\delta}) \end{aligned} \tag{10.2}$$

where $\Sigma^*_{q_0r=q}$ indicates a sum over q_0 with

$$q_0 r = q_1 \quad q_0 = q_1 q_2^2 \dots q_k^k, \quad q_1 \dots q_k \leqslant P^\delta, \quad q_1\dots q_k \in \mathcal{A}(P,R), \quad (r, q_1^{k-1}\dots q_{k-1}) = 1, \tag{10.3}$$

$q_1 \ldots q_k$ squarefree, and

$$T(\gamma, Q, R, r) = \sum_{\substack{y \in \mathscr{A}(Q,R) \\ (r,y)=1}} e(\gamma y^k). \tag{10.4}$$

The next lemma provides a convenient method of factorising the elements of $\mathscr{A}(Q, R)$.

LEMMA 10.1. *Suppose that* $2 \leq R \leq M < y \leq Q$ *and* $y \in \mathscr{A}(Q, R)$. *Then there is a unique triple* (p, u, v) *with*

(i) $y = uv$,

(ii) $u \in \mathscr{A}(Q/v, p)$,

(iii) $M < v \leq Mp$,

(iv) $p | v$,

(v) $p' | v \Rightarrow p \leq p' \leq R$.

Proof. We first show the existence of the triple (p, u, v). Write $y = p_1 \ldots p_s$ where $R \geq p_1 \geq p_2 \geq \ldots$ and let

$$d_j = \prod_{i \leq j} p_i.$$

Then $1 = d_0 < d_1 < \ldots < d_s = y$ and, since $y > R$ and $y \in \mathscr{A}(Q, R)$ we have $s \geq 2$. Since $1 < M < y$ there is a t such that $d_t \leq M < d_{t+1}$. Since $M \geq R$ we have $t \geq 1$, and since $M < y$ we have $t < s$. Therefore $M < d_{t+1} = d_t p_{t+1} \leq M p_{t+1}$. We take $p = p_{t+1}$, $v = d_{t+1}$, $u = y/v$. Clearly each of (i), …, (v) is satisfied.

Now we show the uniqueness of the triple (p, u, v). Suppose that there is another, say (p', u', v'). Without loss of generality we may suppose that either $p' < p$, or $p' = p$ and $v' > v$. Let w and w' denote the largest divisors of y which have all their prime factors exceeding p and p' respectively. Then $v = p^h w$ and $v' = (p')^{h'} w'$ where $h \geq 1$ and $h' \geq 1$.

If $p' < p$, then $v | w'$ so that

$$v' \geq p'v > p'M. \tag{10.5}$$

If $p' = p$, then $w = w'$. Therefore, as $v' > v$, we have $h' > h$. Therefore (10.5) holds in this case also.

Clearly (10.5) contradicts the definition of (p', u', v').

We now apply Lemma 10.1 to T. Note that for any triple (p, u, v) satisfying (ii)–(v) we have $uv \in \mathscr{A}(Q, R)$ and $M < uv \leqslant Q$. Thus, there is a bijection between the y and the (p, u, v). Hence, by (10.4),

$$T(\gamma, Q, R, r) = \sum_{\substack{M<y\leqslant Q \\ y \in \mathscr{A}(Q,R) \\ (r,y)=1}} e(\gamma y^k) + O(M) = \sum_{\substack{p\leqslant R \\ p \nmid r}} U(\gamma, Q, M, R, r, p) + O(M) \tag{10.6}$$

where

$$U(\gamma, Q, M, R, r, p) = \sum_{\substack{v \in \mathscr{B}(M,p,R) \\ (v,r)=1}} \sum_{\substack{u \in \mathscr{A}(Q/v,p) \\ (u,r)=1}} e(\gamma u^k v^k) \tag{10.7}$$

and

$$\mathscr{B}(M, p, R) = \{v\colon M < v \leqslant Mp,\ p|v,\ p'|v \Rightarrow p \leqslant p' \leqslant R\}. \tag{10.8}$$

For $v > M$ we have

$$\sum_{\substack{u \in \mathscr{A}(Q/v,p) \\ (u,r)=1}} e(\gamma u^k v^k) = \int_0^1 \left(\sum_{\substack{u \in \mathscr{A}(Q/M,p) \\ (u,r)=1}} e(\gamma u^k v^k + \theta u) \right) \left(\sum_{x \leqslant Q/v} e(-\theta x) \right) d\theta.$$

Therefore

$$\begin{aligned} U(\gamma, Q, M, R, r, p) &\ll \int_0^1 V(\gamma, Q, M, R, r, p, \theta) \min\left(\frac{Q}{M}, \frac{1}{\|\theta\|}\right) d\theta \\ &\ll (\log Q) \sup_\theta V(\gamma, Q, M, R, r, p, \theta) \end{aligned} \tag{10.9}$$

where

$$V(\gamma, Q, M, R, r, p, \theta) = \sum_{\substack{M<v\leqslant MR \\ (v,r)=1}} \left| \sum_{\substack{u \in \mathscr{A}(Q/M,p) \\ (u,r)=1}} e(\theta u + \gamma v^k u^k) \right|. \tag{10.10}$$

Now we take

$$\gamma = \alpha q_1^k \dots q_k^k, \quad Q = P/(q_1 \dots q_k), \quad M = P^{1/2} R^{-1} (2 q_1^{k-1} q_2^{k-2} \dots q_{k-1})^{-1/k}. \tag{10.11}$$

Note that γ and Q agree with the choices forced upon us when we substitute (10.4) into (10.2).

By (10.2), (10.6) and (10.9),

$$S(\alpha) \ll P^{1-\delta+\varepsilon} + (\log P) \sum_{q_0 r = q}^{*} \sum_{p \leqslant R} \sup_{\theta} V(\gamma, Q, M, R, r, p, \theta). \tag{10.12}$$

When $(h, r)=1$, the number J of solutions of the congruence $x^k \equiv h \pmod r$ satisfies $J \ll r^{\varepsilon}$. Hence there is an $L \ll r^{\varepsilon}$ such that the v with $M < v \leqslant MR$ and $(v, r)=1$ can be divided into L classes $\mathscr{V}_1, \ldots, \mathscr{V}_L$ such that for two distinct elements v_1, v_2 in a given class $\mathscr{V}_j$ we have $v_1^k \equiv v_2^k \pmod r$ if and only if $v_1 \equiv v_2 \pmod r$. Therefore, by (10.10) and Hölder's inequality,

$$V(\gamma, Q, M, R, r, p, \theta)^{2s} \ll P^{\varepsilon}(MR)^{2s-1} \max_j \sum_{v \in \mathscr{V}_j} \left| \sum_{y \leqslant Y} b_y e(\gamma v^k y) \right|^2 \tag{10.13}$$

where

$$Y = s(Q/M)^k \tag{10.14}$$

and $|b_y| \leqslant c_y$ with c_y being the number of solutions of

$$u_1^k + \ldots + u_s^k = y$$

with $u_i \in \mathscr{A}(Q/M, p)$.

By (10.3),

$$\frac{a}{q} q_1^k \ldots q_k^k = a q_1^{k-1} q_2^{k-2} \ldots q_{k-1}/r = a'/r,$$

say, with $(a', r)=1$. Thus, by (10.1), (10.3) and (10.11),

$$\left| \gamma - \frac{a'}{r} \right| \leqslant q_1^{k-1} q_2^{k-2} \ldots q_{k-1}/(r P^{k/2}) = 1/(2r M^k R^k). \tag{10.15}$$

Hence, if $v_1, v_2 \in \mathscr{V}_j$ and $v_1 \not\equiv v_2 \pmod r$, then we have

$$\begin{aligned} \|\gamma(v_1^k - v_2^k)\| &\geqslant \left\| \frac{a'}{r}(v_1^k - v_2^k) \right\| - \frac{1}{2} r^{-1}(MR)^{-k}(MR)^k \\ &\geqslant \frac{1}{2} r^{-1}. \end{aligned}$$

When $r>MR$ the elements of $\mathcal{V}_j$ are distinct modulo r. Thus for $v \in \mathcal{V}_j$ the γv^k are spaced at least $\frac{1}{2}r^{-1}$ apart modulo 1. Therefore, by the large sieve inequality (see, for example, § 27, Theorem 2 of [D4]),

$$\sum_{v \in \mathcal{V}_j} \left| \sum_{y \in Y} b_y e(\gamma v^k y) \right|^2 \ll (Y+r)\, S_s(Q/M, R). \tag{10.16}$$

When $r \leqslant MR$ we have to consider what happens when $v_1 \equiv v_2 \pmod r$ but $v_1 \neq v_2$. Then, by (10.15),

$$\|\gamma(v_1^k - v_2^k)\| = \left\| \left(\gamma - \frac{a'}{r}\right)(v_1^k - v_2^k) \right\| = \left|\gamma - \frac{a'}{r}\right| |v_1^k - v_2^k|.$$

Since $v_1 - v_2$ is a non-zero multiple of r and $v_1 > M$, $v_2 > M$ we have

$$\|\gamma(v_1^k - v_2^k)\| \geqslant \left|\gamma - \frac{a'}{r}\right| M^{k-1} r.$$

Now, by (10.11) and (10.3)

$$q = rq_1q_2^2 \dots q_k^k \leqslant MRq_1q_2^2 \dots q_k^k \leqslant P^{\frac{1}{2}}(q_1 \dots q_k)^k \leqslant P^{\frac{1}{2}+k\delta}.$$

Thus, by the definition of $\mathfrak{m}$, we have $|\alpha - a/q| > q^{-1}P^{(1/2)+k\delta-k}$. Therefore

$$\left|\gamma - \frac{a'}{r}\right| = \left|\alpha - \frac{a}{q}\right| q_1^k \dots q_k^k > P^{\frac{1}{2}+k\delta-k} \frac{q_1^k \dots q_k^k}{rq_1q_2^2 \dots q_k^k}$$

so that, by (10.11),

$$\|\gamma(v_1^k - v_2^k)\| > \frac{1}{2} P^{k\delta - \frac{1}{2}k} R^{1-k}.$$

Thus in this case, for $v \in \mathcal{V}_j$ the γv^k are spaced at least

$$\frac{1}{2}\min(r^{-1}, P^{k\delta - \frac{1}{2}k} R^{1-k})$$

apart modulo 1. Therefore, by the large sieve and (10.16), in either case

$$\sum_{v \in \mathcal{V}_j} \left| \sum_{y \leqslant Y} b_y e(\gamma v^k y) \right|^2 \ll (Y + r + P^{\frac{1}{2}k - k\delta} R^{k-1})\, S_s(Q/M, R).$$

By (10.14) and (10.11), $Y=s(Q/M)^k=2sP^{k/2}R^k/(q_1q_2^2\ldots q_k^k)$, by (10.1) and (10.3), $r\leqslant P^{k/2}/(q_1q_2^2\ldots q_k^k)$, and, by (10.3), $q_1q_2^2\ldots q_k^k\leqslant P^{k\delta}$. Thus, by (10.13) and Theorem 7.1,

$$V(\gamma,Q,M,R,r,p,\theta)^{2s}\ll P^{\varepsilon}(MR)^{2s-1}(Q/M)^k(Q/M)^{2s-k+\sigma}$$

where

$$\sigma=(k-2)\left(1-\frac{1}{k}\right)^{s-2}.$$

Hence, by (10.12) and (10.11),

$$S(\alpha)\ll P^{1-\delta+\varepsilon}+\sum_{q_0r=q}^{*}\sum_{p\leqslant R}P^{\varepsilon}R^{1-\frac{1}{2s}}Q(Q/M)^{\frac{\sigma}{2s}}M^{-\frac{1}{2s}}$$

$$\ll P^{1-\delta+\varepsilon}+R^{2+\frac{\sigma}{2s}}P^{1-\frac{1-\sigma}{4s}+\varepsilon}\sum_{q_0r=q}(q_1\ldots q_k)^{\frac{1}{2s}-1}$$

and the theorem follows.

References

[BM] Balasubramanian, R. & Mozzochi, C. J., An improved upper bound for $G(k)$ in Waring's problem for relatively small k. *Acta Arith.*, 43 (1984), 283–285.

[B1] de Bruijn, N. G., The asymptotic behaviour of a function occurring in the theory of primes. *J. Indian Math. Soc. (N.S.)*, 15 (1951), 25–32.

[B2] — On the number of positive integers $\leqslant x$ and free of prime factors $>y$. *Nederl. Acad. Wetensch. Proc. Ser. A.*, 54 (1951), 50–60.

[C] Chen, Jing-Run, On Waring's problem for nth powers. *Acta Math. Sinica*, 8 (1958), 253–257; translated in *Chin. Math. Acta*, 8 (1966), 849–853.

[D1] Davenport, H., On Waring's problem for fourth powers. *Ann. of Math.*, 40 (1939), 731–747.

[D2] — On sums of positive integral kth powers. *Amer. J. Math.*, 64 (1942), 189–198.

[D3] — *The collected works of Harold Davenport, vol.* III. Ed. B. J. Birch, H. Halberstam & C. A. Rogers, Academic Press, 1977.

[D4] — *Multiplicative number theory*. Springer-Verlag, second edition, 1980.

[DE] Davenport, H. & Erdős, P., On sums of positive integral kth powers. *Ann. of Math.*, 40 (1939), 533–536.

[DL] Davenport, H. & Lewis, D. J., Homogeneous additive equations. *Proc. Roy. Soc. London Ser. A*, 274 (1963), 443–460.

[E] Estermann, T., Einige Sätze über quadratfreie Zahlen. *Math. Ann.*, 105 (1931), 653–662.

[HR] Halberstam, H. & Richert, H.-E., *Sieve methods*. Academic Press, 1974.

[HL] Hardy, G. H. & Littlewood, J. E., *Collected papers of G. H. Hardy, including joint papers with J. E. Littlewood and others*. Ed. by a committee appointed by the London Mathematical Society, vol. I, Clarendon Press, 1966.

[H] Hooley, C., On Waring's problem. *Acta Math.*, 57 (1986), 49–97.

[K] Karatsuba, A. A., On the function $G(n)$ in Waring's problem. *Izv. Akad. Nauk SSSR*, 49 (1985), 935–947, 1119.
[L] Linnik, Ju. V., On the representation of large numbers as sums of seven cubes. *Dokl. Akad. Nauk SSSR*, 35 (1942), 162 and *Mat. Sb.*, 12 (1943), 218–224.
[T1] Thanigasalam, K., On Waring's problem. *Acta Arith.*, 38 (1980), 141–155.
[T2] — Some new estimates for $G(k)$ in Waring's problem. *Acta Arith.*, 42 (1982), 73–78.
[T3] — Improvement on Davenport's iterative method and new results in additive number theory, I & II, proof that $G(5)\leq 22$. *Acta Arith.*, 46 (1985), 1–31 and 91–112.
[Va1] Vaughan, R. C., Homogeneous additive equations and Waring's problem. *Acta Arith.*, 33 (1977), 231–253.
[Va2] — *The Hardy-Littlewood method*. Cambridge University Press, 1981.
[Va3] — Sums of three positive cubes, *Bull. London Math. Soc.*, 17 (1985), 17–20.
[Va4] — On Waring's problem for smaller exponents. *Proc. London Math. Soc. (3)*, 52 (1986), 445–63.
[Va5] — On Waring's problem for sixth powers. *J. London Math. Soc. (2)*, 33 (1986), 227–236.
[Va6] — On Waring's problem for cubes. *J. Reine Angew. Math.*, 365 (1986), 122–170.
[Va7] — On Waring's problem for smaller exponents II. *Mathematika*, 33 (1986), 6–22.
[Vi1] Vinogradov, I. M., The method of trigonometrical sums in the theory of numbers. *Trudy Mat. Inst. Steklov*, 23 (1947), 1–109.
[Vi2] — On an upper bound for $G(n)$. *Izv. Akad. Nauk SSSR*, 23 (1959), 637–642.
[Vi3] — *Selected works*. Springer-Verlag.
[W] Watson, G. L., A proof of the seven cube theorem. *J. London Math. Soc.*, 26 (1951), 153–156.

Received June 24, 1987
Received in revised form June 11, 1988