ON EXPONENTIAL SUMS OVER SMOOTH NUMBERS

Trevor D. Wooley

1. Introduction

This paper is concerned with the theory and applications of exponential sums over smooth numbers. Despite the numerous applications stemming from suitable estimates for these exponential sums, thus far little attention has been paid to any but the simplest cases (see, for example, [19, 20, 23]). Our primary objective is the development of a method for estimating mean values of exponential sums over smooth numbers, such mean values being fundamental to subsequent applications. Having established such a method, we develop estimates of use in applications of such exponential sums inside the fabric of the Hardy-Littlewood method. For the purposes of illustrating the power of our new estimates, we draw corollaries concerning the distribution of the fractional parts of polynomials, and for Waring's problem with polynomial summands. There are also consequences of our methods for problems involving the global solubility of simultaneous additive equations, but we defer an account of such developments to a later occasion (see [30]). These applications by no means exhaust the available supply. Our estimates will also be useful in considering problems concerning simultaneous small values of additive forms (see, for example, [18], Chapter 11), and the simultaneous distribution modulo 1 of additive forms (see [6]).

In order to describe our conclusions, we shall require some notation. When P and R are positive integers, let $\mathcal{A}(P,R)$ denote the set of R-smooth numbers up to P, that is,

$$\mathcal{A}(P,R)=\{n\in\mathbb{Z}\cap[1,P]\ :\ p \text{ prime},\ p|n\Rightarrow p\le R\}. \tag{1.1}$$

Consider a fixed t-tuple, $\mathbf{k}=(k_1,\dots,k_t)$, of positive integers satisfying

$$1\le k_t<k_{t-1}<\dots<k_1, \tag{1.2}$$

and define the exponential sum $f_{\mathbf{k}}(\boldsymbol{\alpha};P,R)$ by

$$f_{\mathbf{k}}(\boldsymbol{\alpha};P,R)=\sum_{x\in\mathcal{A}(P,R)}e(\alpha_1x^{k_1}+\dots+\alpha_tx^{k_t}), \tag{1.3}$$

where $e(z)$ denotes $e^{2\pi iz}$. We define $S_s(P,R)=S_s^{(\mathbf{k})}(P,R)$ by

$$S_s^{(\mathbf{k})}(P,R)=\int_{\mathbb{T}^t}|f_{\mathbf{k}}(\boldsymbol{\alpha};P,R)|^{2s}d\boldsymbol{\alpha}, \tag{1.4}$$

where $\mathbb{T}^t$ denotes the t-dimensional unit cube. We note that by orthogonality, $S_s(P,R)$ is equal to the number of solutions of the system of diophantine equations

$$\sum_{i=1}^{s}(x_i^{k_j}-y_i^{k_j})=0\quad(1\le j\le t), \tag{1.5}$$

with $x_i,y_i\in\mathcal{A}(P,R)$ $(1\le i\le s)$.

Research initially supported in part by NSF grant DMS-8610730 at the Institute for Advanced Study, a Rackham Fellowship from the University of Michigan, and a SERC Visiting Research Fellowship at Imperial College, London. This paper was completed with the aid of NSF grant DMS-9303505, an Alfred P. Sloan Research Fellowship and a Fellowship from the David and Lucile Packard Foundation.

Our estimates for $S_s(P,R)$ are established through iterative procedures similar to those of Vaughan and Wooley [19, 20, 23]. In order to explain such methods, it is convenient to describe some notation with which to discuss bounds for the mean values $S_s(P,R)$. We shall say that an exponent $\lambda_s=\lambda_{s,\mathbf{k}}$ is *permissible* whenever the exponent has the property that for each $\varepsilon>0$, there exists a positive number $\eta=\eta(\varepsilon,s,\mathbf{k})$ such that whenever $R\le P^\eta$, one has

$$S_s^{(\mathbf{k})}(P,R)\ll_{\varepsilon,s,\mathbf{k}} P^{\lambda_{s,\mathbf{k}}+\varepsilon}. \tag{1.6}$$

Permissible exponents certainly exist, since for each s the estimate $S_s^{(\mathbf{k})}(P,R)\le P^{2s}$ is trivial. In our applications we take $R=P^\eta$ with η small and positive. In such circumstances the bound $\text{card}(\mathcal{A}(P,R))\gg_\eta P$ leads via a standard argument (see, for example, [25]) to the lower bound

$$S_s^{(\mathbf{k})}(P,R)\gg_{s,\mathbf{k},\eta} P^s+P^{2s-\sum_{i=1}^t k_i}, \tag{1.7}$$

whence whenever the exponent $\lambda_{s,\mathbf{k}}$ is permissible, one has $\lambda_{s,\mathbf{k}}\ge 2s-\sum_{i=1}^t k_i$. We note that by considering products of local densities, for $s>k_1+\dots+k_t$ one expects that the exponent $\lambda_{s,\mathbf{k}}=2s-\sum_{i=1}^t k_i$ should be permissible.

In its most basic form, the efficient differencing process which we develop in §§2, 3 and 4, leads to very simple, yet useful bounds for permissible exponents. In §5 we establish the following theorem.

Theorem 1. *Let $k_1,\dots,k_t$ be integers satisfying* (1.2). *Suppose that r is an integer with $1<r\le t+1$, and that s is a positive integer with $s\equiv r \pmod t$. Then the exponent*

$$\lambda_{s,\mathbf{k}}=2s-\sum_{i=1}^t k_i+\Delta_s$$

is permissible, where

$$\Delta_s=\left(\sum_{i=1}^t k_i-r\right)(1-1/k_1)^{(s-r)/t}.$$

When $t=1$, Theorem 1 provides the same conclusion as Vaughan [19], Theorem 7.1. We note that Karatsuba has stated a similar theorem ([13], Theorem 1) which provides estimates for $S_s(P,R)$, with a modest condition on s, in the more restricted range of R satisfying $\log R=o(\log P)$. Karatsuba provides no details of the proof of his theorem, but notes that his argument emulates Linnik's so-called "p-adic" approach to Vinogradov's mean value theorem (see [14]). In common with previous work on mean values of exponential sums over smooth numbers, our proof of Theorem 1 is also motivated by Linnik's p-adic method. However, serious obstacles must be negotiated in order to obtain the full strength of Theorem 1 above.

In §6 we exploit the full power of our repeated efficient differencing method, thereby improving substantially the estimates provided by Theorem 1 (and also, in consequence, the estimates of Karatsuba [13], Theorem 1). Our conclusions are somewhat complicated to describe in their most precise form, and so we defer their enunciation to §6. Fortunately, however, it is possible to obtain simple estimates of considerable utility, the first of which we describe in Theorem 2 below.

Theorem 2. *Let $k_1,\dots,k_t$ be integers satisfying* (1.2). *Write*

$$s_0=\tfrac12 tk_1\left(\log(tk_1)-2\log\log k_1\right),$$

and when s is a positive integer, define the exponent Δ_s by

$$\Delta_s=\begin{cases} tk_1e^{2-2s/(tk_1)}, & \text{when } 1\le s\le s_0,\\ e^3(\log k_1)^2\,(1-1/k_1)^{(s-s_0)/t}, & \text{when } s>s_0.\end{cases} \tag{1.8}$$

Then there exists an absolute constant K_1 such that whenever $k_1\ge K_1$, the exponent $\lambda_{s,\mathbf{k}}=2s-\sum_{i=1}^t k_i+\Delta_s$ is permissible.

Observe that Theorem 2 yields permissible exponents $\lambda_s=2s-\sum_{i=1}^t k_i+\Delta_s$ with Δ_s behaving roughly like $tk_1e^{-2s/(tk_1)}$. Meanwhile Theorem 1 yields a similar conclusion with Δ_s behaving in most circumstances like $tk_1e^{-s/(tk_1)}$, and consequently Theorem 2 is about twice as powerful as Theorem 1 in applications. We note that while the factor e^3 occurring in definition (1.8) can be somewhat reduced with greater effort, such an improvement does not have significant consequences.

When t is no larger than about $(\log(k_1\dots k_t))^{1/2}$ and s is sufficiently large, an alternative strategy leads us in §6 to a conclusion which usually provides estimates for permissible exponents λ_s superior to those stemming from Theorem 2.

Theorem 3. *Let $t \geq 2$, and let $k_1, \ldots, k_t$ be integers satisfying* (1.2). *Write*

$$s_1 = \left[\tfrac{1}{2}k_1\left(\log(k_1k_2\ldots k_t) + 3t^2\right)\right], \tag{1.9}$$

and when s is an integer exceeding s_1 with $s \equiv s_1 \pmod{t}$, define the exponent Δ_s by

$$\Delta_s = (\log k_1)^2(1 - 1/k_1)^{(s-s_1)/t}. \tag{1.10}$$

Then there exists an absolute constant K_2 such that whenever $k_1 \geq K_2$, the exponent $\lambda_{s,\mathbf{k}} = 2s - \sum_{i=1}^t k_i + \Delta_s$ is permissible.

Theorem 3 is particularly effective when only a small number of the k_i are large, since in such circumstances $\log(k_1 \ldots k_t)$ is substantially smaller than $t \log k_1$.

In the second part of this paper we consider several applications of our new mean value estimates. In §7 we establish upper bounds for smooth Weyl sums of use on the minor arcs in applications of the Hardy-Littlewood method. In Theorem 4 we provide such bounds for multi-dimensional exponential sums.

Theorem 4. *When $k_1, \ldots, k_t$ are integers satisfying* (1.2), *and μ is a positive real number, denote by $\mathfrak{m}_\mu$ the set of $\boldsymbol{\alpha} \in \mathbb{R}^t$ such that whenever $\mathbf{a} \in \mathbb{Z}^t$, $q \in \mathbb{N}$,*

$$(a_1, \ldots, a_t, q) = 1 \quad \textit{and} \quad |\alpha_i - a_i/q| \leq q^{-1}P^{\mu - k_i}R^{t-1} \quad (1 \leq i \leq t), \tag{1.11}$$

then one has $q > P^\mu R^t$. Suppose that t is an integer exceeding 1, and that λ is a real number with $0 < \lambda \leq 1/t$. Suppose further that the numbers Δ_s ($s \in \mathbb{N}$) have the property that for each s the exponent $\lambda_{s,\mathbf{k}} = 2s - \sum_{i=1}^t k_i + \Delta_s$ is permissible. Then for each $\varepsilon > 0$, there exists a positive number $\eta = \eta(\varepsilon, \lambda, \mathbf{k})$ such that whenever $R \leq P^\eta$, one has

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

where

$$\sigma(\mathbf{k}; \lambda) = \max_{2s \geq k_1+1} \frac{\lambda - (1-\lambda)\Delta_s}{2s}. \tag{1.12}$$

When the argument of the exponential sum $f_{\mathbf{k}}(\boldsymbol{\alpha}; P, R)$ takes the shape $\alpha\psi(x)$, with $\psi(x) \in \mathbb{Z}[x]$, one can obtain a sharper conclusion. We define some further notation at this point for the sake of convenience. When $a_i \in \mathbb{Z}$ $(1 \leq i \leq t)$ we write $\psi_{\mathbf{k}}(x; \mathbf{a}) = a_1x^{k_1} + \cdots + a_tx^{k_t}$, and define

$$g_{\mathbf{k}}(\alpha; P, R) = \sum_{x \in \mathcal{A}(P,R)} e(\alpha\psi_{\mathbf{k}}(x; \mathbf{a})). \tag{1.13}$$

Theorem 5. *Suppose that $k_1, \ldots, k_t$ are integers satisfying* (1.2), *and when μ is a positive real number, denote by $\mathfrak{n}_\mu$ the set of $\alpha \in \mathbb{R}$ such that whenever $a \in \mathbb{Z}$, $q \in \mathbb{N}$, $(a, q) = 1$ and $|\alpha - a/q| \leq q^{-1}P^{\mu - k_1}$, then one has $q > |a_1|P^\mu R$. Suppose that λ is a positive real number with $0 < \lambda \leq \frac{1}{2}$. Then with $\sigma(\mathbf{k}; \lambda)$ defined by* (1.12), *for each $\varepsilon > 0$, there exists a positive number $\eta = \eta(\varepsilon, \lambda, \mathbf{k})$ such that whenever $R \leq P^\eta$,*

$$\sup_{\alpha \in \mathfrak{n}_\lambda} |g_{\mathbf{k}}(\alpha; P, R)| \ll P^{1-\sigma(\mathbf{k};\lambda)+\varepsilon}.$$

There are immediate applications of Theorems 4 and 5 to problems concerning localised estimates for the fractional parts of polynomials. Thus, in Theorem 6 we provide estimates for $\min_{1 \leq n \leq N} \|f(n)\|$, where $f(n)$ is a polynomial and $\|x\|$ denotes $\min_{y \in \mathbb{Z}} |x - y|$.

Theorem 6. *Let $k_1, \ldots, k_t$ be integers satisfying* (1.2), *and let $\boldsymbol{\alpha} \in \mathbb{R}^t$. Suppose that the numbers Δ_s ($s \in \mathbb{N}$) have the property that for each s the exponent $\lambda_{s,\mathbf{k}} = 2s - \sum_{i=1}^t k_i + \Delta_s$ is permissible, and define $\tau(\mathbf{k})$ by*

$$\tau(\mathbf{k}) = \max_{2s \geq k_1+1} \frac{1 - (t-1)\Delta_s}{2st + 1 + \Delta_s}.$$

Then for each $\varepsilon > 0$ there is a real number $N_0 = N_0(\mathbf{k}, \varepsilon, t)$ such that whenever $N > N_0$,

$$\min_{1\le n\le N} \|\alpha_1 n^{k_1} + \dots + \alpha_t n^{k_t}\| < N^{\varepsilon-\tau(\mathbf{k})}.$$

Moreover, if the α_i are in rational ratio, so that $\boldsymbol{\alpha} = \alpha\mathbf{a}$ for some rational t-tuple $\mathbf{a}$ and real number α, then for each $\varepsilon > 0$ there is a real number $N_1 = N_1(\mathbf{k}, \varepsilon, t)$ such that whenever $N > N_1$,

$$\min_{1\le n\le N} \|\alpha_1 n^{k_1} + \dots + \alpha_t n^{k_t}\| < N^{\varepsilon-\sigma(\mathbf{k};1/2)},$$

where $\sigma(\mathbf{k}; \lambda)$ is defined as in (1.12).

It seems worthwhile to record the bounds stemming from Theorems 4, 5 and 6 for large k_1 in the following theorem.

Theorem 7. *Suppose that $k_1, \dots, k_t$ are integers satisfying* (1.2)*, and that k_1 is large. Define $\sigma_i(\mathbf{k})$ $(i = 1, 2)$ by*

$$\sigma_1(\mathbf{k})^{-1} = tk_1 \min\left\{t(\log k_1 + 3\log t), 3t^2 + 6t\log\log k_1 + \log(k_1 \dots k_t)\right\},$$

and

$$\sigma_2(\mathbf{k})^{-1} = 2k_1 \min\left\{t(\log k_1 + \log t), 3t^2 + 6t\log\log k_1 + \log(k_1 \dots k_t)\right\}.$$

(i) Defining $\mathfrak{m}_1$ and $\mathfrak{n}_{1/2}$ as in the statements of Theorems 4 *and* 5*, for each $\varepsilon > 0$, there exists a positive number $\eta = \eta(\varepsilon, \mathbf{k})$ such that whenever $R \le P^\eta$,*

$$\sup_{\boldsymbol{\alpha}\in\mathfrak{m}_1} |f_{\mathbf{k}}(\boldsymbol{\alpha}; P, R)| \ll P^{1-(1+o(1))\sigma_1(\mathbf{k})+\varepsilon}$$

and

$$\sup_{\alpha\in\mathfrak{n}_{1/2}} |g_{\mathbf{k}}(\alpha; P, R)| \ll P^{1-(1+o(1))\sigma_2(\mathbf{k})+\varepsilon}.$$

(ii) Let $\boldsymbol{\alpha} \in \mathbb{R}^t$. Then for each $\varepsilon > 0$ there is a real number $N_0 = N_0(\mathbf{k}, \varepsilon, t)$ such that whenever $N > N_0$,

$$\min_{1\le n\le N} \|\alpha_1 n^{k_1} + \dots + \alpha_t n^{k_t}\| \ll N^{\varepsilon-(1+o(1))\sigma_1(\mathbf{k})}.$$

Moreover if the α_i are in rational ratio then $\sigma_1(\mathbf{k})$ may be replaced by $\sigma_2(\mathbf{k})$ in the latter conclusion.

For comparison, Corollary 1.3 of Wooley [24] improves on Theorem 4.5 of Baker [1] to provide the conclusion

$$\min_{1\le n\le N} \|\alpha_1 n + \dots + \alpha_k n^k\| \ll N^{\varepsilon-\sigma_3(k)},$$

with $\sigma_3(k)^{-1} \sim 4k^2 \log k$, and Theorem 1.2 of Wooley [28] yields

$$\min_{1\le n\le N} \|\alpha n^k\| \ll N^{\varepsilon-\sigma_4(k)},$$

with $\sigma_4(k)^{-1} \sim k\log k$ (similar conclusions hold for the corresponding estimates for Weyl sums, and smooth Weyl sums, in the respective cases). It follows, in particular, that Theorem 7 improves on the estimates available hitherto for t in the range $2 \le t \le k_1^{1/2}$. In order to discuss these improvements further it is convenient to introduce some nomenclature. We will describe a polynomial $f(x; \boldsymbol{\alpha}) = \sum_{i=1}^t \alpha_i x^{k_i}$ as being of *weight* t if the total number of non-zero coefficients is t. Thus a non-trivial polynomial with non-zero degree k has weight between 1 and k (note that we assume the constant term of the polynomial to be zero for the purposes of these deliberations). Further, we will refer to the polynomial f as being *d-lite* if $k_1, \dots, k_t$ satisfy (1.2) and $d > k_2$. Thus every polynomial of degree k is k-lite, and the monomial x^k is 1-lite. We note that in those situations where $\psi(x; \boldsymbol{\alpha})$ has weight t with $t = o(\sqrt{\log(k_1 \dots k_t)})$, one has $\sigma_1(\mathbf{k})^{-1} = (1+o(1))tk_1 \log(k_1 \dots k_t)$ and $\sigma_2(\mathbf{k})^{-1} = (2+o(1))k_1 \log(k_1 \dots k_t)$, and so when the k_i are suitably distributed, Theorem 7 yields substantial improvements over the previous state of knowledge. For example, when $\psi(x; \boldsymbol{\alpha})$ has weight t with $t = o(\sqrt{\log k_1})$, and is d-lite with $\log d = O(t)$, then a simple calculation reveals that $\sigma_2(\mathbf{k})^{-1} = (2 + o(1))k_1 \log k_1$, and thus the conclusions of Theorem 7 are of similar strength to results obtained hitherto for the monomial αx^{k_1}.

In §8 we turn our attention to unlocalised estimates for fractional parts of polynomials. We refer the reader to §8 for the details of the exponential sum estimates required in our argument, and instead record the conclusions of that argument in Theorem 8 below.

Theorem 8. *Let $k_1,\dots,k_t$ be integers satisfying* (1.2)*, and let $\boldsymbol{\alpha}$ be a real t-tuple with the property that α_i is irrational for $1 \le i \le t$. Suppose that the numbers Δ_s $(s \in \mathbb{N})$ have the property that for each s the exponent $\lambda_{s,\mathbf{k}} = 2s - \sum_{i=1}^t k_i + \Delta_s$ is permissible. Then for each $\varepsilon > 0$ there are infinitely many $n \in \mathbb{N}$ with*

$$\|\alpha_1 n^{k_1} + \cdots + \alpha_t n^{k_t}\| \le n^{\varepsilon - \alpha(\mathbf{k})}, \tag{1.14}$$

where

$$\alpha(\mathbf{k}) = \max_{s\in\mathbb{N}} \frac{k_1 - 2\Delta_s}{4s^2}.$$

Moreover, if the α_i are in rational ratio, so that $\boldsymbol{\alpha} = \alpha\mathbf{a}$ for some rational t-tuple $\mathbf{a}$ and real number α, then for each $\varepsilon > 0$ there are infinitely many $n \in \mathbb{N}$ with

$$\|\alpha_1 n^{k_1} + \cdots + \alpha_t n^{k_t}\| \le n^{\varepsilon - \max\{\alpha(\mathbf{k}),\beta(\mathbf{k})\}},$$

where

$$\beta(\mathbf{k}) = \max_{\substack{s\in\mathbb{N}\\ s\ge t}} \frac{\mathcal{K} - 2\Delta_s}{4s^2},$$

and $\mathcal{K} = \sum_{i=1}^t \max\{0, 2k_i - k_1\}$.

We illustrate the consequences of Theorem 8 with the following corollary.

Corollary 8.1. *When $t > 1$ and k_1 is large, the conclusion* (1.14) *of Theorem* 8 *holds with*

$$\alpha(\mathbf{k})^{-1} = t^2 k_1 (\log t + \log\log t + 7)^2.$$

Moreover, if the polynomial $\alpha_1 x^{k_1} + \cdots + \alpha_t x^{k_t}$ is d-lite with $d = o(\sqrt{k_1})$, then (1.14) *holds with $\alpha(\mathbf{k})^{-1} = (\gamma + o(1))k_1$, where $\gamma = 9.0267256...$ is defined by $\gamma = (\omega + \log\omega - 1)^2/(1 - 2\omega)$, in which ω is the unique positive solution of the equation $\omega + 2 - \omega^{-1} = \log\omega$.*

We note that the second conclusion of Corollary 8.1 is of the same strength as that obtained in [26], Corollary 1 to Theorem 1.1 for the monomial αx^{k_1}. We note also that when the α_i are in rational ratio, and the k_i are of size comparable to k_1, then the second part of Theorem 8 yields an exponent $\beta(\mathbf{k})$ with $\beta(\mathbf{k})^{-1} \asymp tk_1$.

Finally, in §§9 and 10, we investigate the consequences of our new mean value estimates for Waring's problem with polynomial summands, a problem which has experienced little progress since work 40 years ago of Hua [8, 9, 10, 11], Nečaev [15, 16, 17] and Chen [2]. In order to describe our conclusions we require some notation. Let $g(x)$ be a polynomial with rational coefficients taking integral values whenever the argument, x, is an integer. Suppose that the degree of $g(x)$ is k, and that its leading coefficient is the non-zero rational number a. When s is a natural number and n is a positive integer, let $R_{s,g}(n)$ denote the number of representations of n in the form

$$g(x_1) + \cdots + g(x_s) = n, \tag{1.15}$$

with the x_i non-negative integers. Then when s is sufficiently large in terms of k, one may apply the Hardy-Littlewood method to obtain the asymptotic formula

$$R_{s,g}(n) = a^{-s/k} \frac{\Gamma(1+1/k)^s}{\Gamma(s/k)} \mathfrak{S}_{s,g}(n) n^{s/k-1} + o(n^{s/k-1}), \tag{1.16}$$

where $\mathfrak{S}_{s,g}(n)$ denotes the *singular series*, defined by

$$\mathfrak{S}_{s,g}(n) = \sum_{q=1}^{\infty} \sum_{\substack{a=1\\(a,q)=1}}^{q} \left(q^{-1} S_g(q,a)\right)^s e(-an/q), \tag{1.17}$$

where

$$S_g(q,a) = \sum_{r=1}^{q} e\left(ag(r)/q\right). \tag{1.18}$$

When $g(x)=x^k$ the formula (1.16) is known to hold for $s\ge s_0(k)$, where $s_0(k)\sim k^2\log k$ (see [3]). In this case, moreover, (1.16) is genuinely an asymptotic formula, since when $s\ge 4k$ the singular series (1.17) converges absolutely to a positive number bounded away from zero (see [18], Theorem 4.6). Unfortunately the behaviour of the singular series is considerably more complicated for general polynomials. Plainly, if for every integer x the polynomial $g(x)$ is divisible by some integer d exceeding 1, then $\mathfrak{S}_{s,g}(n)$ will be zero when $d\nmid n$. Even those polynomials $g(x)$ for which no such d exists may exhibit complicated behaviour. Consider, for example, the polynomials $H_k(x)$ defined by

$$H_k(x)=2^{k-1}F_k(x)-2^{k-2}F_{k-1}(x)+\dots+(-1)^{k-1}F_1(x),$$

where $F_i(x)=x(x-1)\dots(x-i+1)/i!$ $(1\le i\le k)$. When $k\ge 5$, Hua [11] has shown that for $s<2^k-\frac{1}{2}(1-(-1)^k)$, there is a certain arithmetic progression of integers n for which the equation $H_k(x_1)+\dots+H_k(x_s)=n$ is locally insoluble, whence $\mathfrak{S}_{s,H_k}(n)=0$.

Rather than consider the intricacies of the behaviour of the singular series in this problem, which is hardly the point of the present paper, we refer the reader to Chapters 1–3 of Nečaev [16], and remark only that Hua [11] has shown that when g has degree k and $s\ge(k-1)2^{k+1}$, then $\mathfrak{S}_{s,g}(n)\gg_g 1$. We proceed along simpler lines, defining $\overline{G}(g)$ to be the least number s satisfying the property that given a positive number δ, all sufficiently large numbers n with $\mathfrak{S}_{s,g}(n)>\delta$ are represented in the form (1.15). In §§9 and 10 we apply Theorems 2, 3 and 5 to obtain the following estimates for $\overline{G}(g)$.

Theorem 9. *Let $g(x)$ be a polynomial with rational coefficients taking integral values whenever the argument, x, is an integer. Suppose that g has degree k and weight t. Then*

$$\overline{G}(g)\le 2k(\log k+\log t+\log\log k+O(1)).$$

Suppose further that $g(x)=\sum_{i=1}^t a_ix^{k_i}$, with $k_1,\dots,k_t$ satisfying (1.2), *and that the a_i are non-zero.*

(i) If $t=o\left(\sqrt{\log(k_1\dots k_t)}\right)$, then $\overline{G}(g)\le(1+o(1))k_1\log(k_1\dots k_t)$.

(ii) If g is d-lite with $d=o(\log k_1/\log\log k_1)$, then $\overline{G}(g)\le(1+o(1))k_1\log k_1$.

We note that a standard argument involving the use of diminishing ranges combined with Vinogradov's estimates for exponential sums yields the bound $\overline{G}(g)\le(4+o(1))k\log k$ for any polynomial g of degree k. Meanwhile, in the classical version of Waring's problem, in which $g(x)=x^k$, the best available bound for $\overline{G}(g)$ is $\overline{G}(g)\le k(\log k+\log\log k+2+o(1))$ (see Theorem 1.4 of Wooley [28]). Thus the first bound of Theorem 9 interpolates between the latter two bounds for polynomials with weight a power of k, and the final bound is essentially as strong as the best that can be proved for the special case of a monomial.

This paper has benefitted from conversations with a number of people over the several years it took to complete, and the author gladly thanks them all. The author gratefully acknowledges the many helpful comments of the referee. The first version of this paper was drafted during my first year of graduate studies at Imperial College, London, directed by Professor R. C. Vaughan, and partly supported by a grant from the SERC (now the EPSRC).

Part I. Mean Value Estimates

2. Some preliminary lemmata

Our first goal in this section is an analogue of Linnik's Lemma suitable for our application in §3. The argument we use to establish this result is an elaboration on the treatments of [24], Lemma 2.2 and [27], Lemma 2.1. Fundamental to our argument is the use of Lemma 2.1 below, which is closely related to Bézout's Theorem.

Lemma 2.1. *Let $f_1,\dots,f_d$ be polynomials in $\mathbb{Z}[x_1,\dots,x_d]$ with respective degrees $k_1,\dots,k_d$, and write*

$$J(\mathbf{f};\mathbf{x})=\det\left(\frac{\partial f_j}{\partial x_i}(\mathbf{x})\right)_{1\le i,j\le d}.$$

When p is a prime number, and s is a natural number, let $\mathcal{N}(\mathbf{f};p^s)$ denote the number of solutions of the simultaneous congruences

$$f_j(x_1,\dots,x_d)\equiv 0\pmod{p^s}\quad(1\le j\le d),$$

with $1 \le x_i \le p^s$ $(1 \le i \le d)$ *and* $(J(\mathbf{f};\mathbf{x}),p)=1$. *Then* $\mathcal{N}(\mathbf{f};p^s) \le k_1 \dots k_d$.

Proof. This is Theorem 1 of [29].

In order to describe our analogue of Linnik's Lemma, we shall require some notation. Let r, t and $k_1,\dots,k_t$ be natural numbers satisfying (1.2), and let $f_1,\dots,f_t$ be polynomials in $\mathbb{Z}[x_1,\dots,x_r]$. When d is an integer with $1 \le d \le \min\{t,r\}$, define the Jacobian determinant $J_d(\mathbf{f};\mathbf{x})$ by

$$J_d(\mathbf{f};\mathbf{x}) = \det\left(\frac{\partial f_i}{\partial x_j}(\mathbf{x})\right)_{1\le i,j\le d}. \tag{2.1}$$

When k and m are positive integers, and $\mathbf{u} \in \mathbb{Z}^t$, define $\mathcal{B}^{\mathbf{k}}_{k,d}(m;\mathbf{u};\mathbf{f})$ to be the set of solutions of the simultaneous congruences

$$f_j(x_1,\dots,x_r) \equiv u_j \pmod{m^{k_j}} \quad (1 \le j \le t),$$

with $x_1,\dots,x_r$ distinct modulo m^k, and satisfying the condition $(J_d(\mathbf{f};\mathbf{x}),m)=1$. Finally, define the function $\omega(\mathbf{k};k,r,d)$ by

$$\omega(\mathbf{k};k,r,d) = (r-d)k + \sum_{j\in\mathcal{I}}(k-k_j), \tag{2.2}$$

where $\mathcal{I} = \mathcal{I}(\mathbf{k};k,d)$ denotes the set of those indices j for which $k_j < k$ and $1 \le j \le d$.

Lemma 2.2. *Suppose that* $f_1,\dots,f_t$ *are polynomials in* $\mathbb{Z}[x_1,\dots,x_r]$ *with degrees bounded in terms of* $k_1,\dots,k_t$ *alone. Suppose also that* $1 \le d \le \min\{t,r\}$. *Then*

$$\mathit{card}\left(\mathcal{B}^{\mathbf{k}}_{k,d}(m;\mathbf{u};\mathbf{f})\right) \ll_{\varepsilon,\mathbf{k}} m^{\omega(\mathbf{k};k,r,d)+\varepsilon}.$$

Proof. We start with the trivial observation that

$$\mathrm{card}\left(\mathcal{B}^{\mathbf{k}}_{k,d}(m;\mathbf{u};\mathbf{f})\right) \le \mathrm{card}\left(\mathcal{B}^{\mathbf{k}^*}_{k,d}(m;\mathbf{u};\mathbf{f})\right),$$

where $k_i^* = \min\{k,k_i\}$ $(1 \le i \le t)$. It is well-known that the number of prime divisors of an integer m with $m \ge 3$ is $O(\log m/\log\log m)$ (see, for example, §22.10 of Hardy and Wright [4]). Thus, by the Chinese Remainder Theorem, it suffices to show that for any prime p, and each positive integer h,

$$\mathrm{card}\left(\mathcal{B}^{\mathbf{k}^*}_{k,d}(p^h;\mathbf{u};\mathbf{f})\right) \ll_{\mathbf{k}} \left(p^h\right)^{\omega(\mathbf{k};k,r,d)}. \tag{2.3}$$

When v is a positive integer, define $h_j(v)$ to be v when $1 \le j \le d$, and to be 1 otherwise. We start by considering $\mathcal{C}^s_v(p;\mathbf{a};\mathbf{b})$, which for $d \le s \le r$ we define to be the number of solutions $(z_1,\dots,z_s)$ distinct modulo p^v of the system of congruences

$$f_j(z_1,\dots,z_s,b_1,\dots,b_{r-s}) \equiv a_j \pmod{p^{h_j(v)}} \quad (1 \le j \le t), \tag{2.4}$$

with $(z_1,\dots,z_s)$ satisfying the condition $(J_d(\mathbf{f};\mathbf{z}),p)=1$. Notice that $\mathcal{C}^s_v(p;\mathbf{a};\mathbf{b})$ is independent of $\mathbf{b}$ when $s=r$. In such circumstances we abbreviate $\mathcal{C}^r_v(p;\mathbf{a};\mathbf{b})$ to $\mathcal{C}^r_v(p;\mathbf{a})$. We bound $\mathcal{B}^{\mathbf{k}^*}_{k,d}$ through the inequality

$$\mathrm{card}\left(\mathcal{B}^{\mathbf{k}^*}_{k,d}(p^h;\mathbf{u};\mathbf{f})\right) \le \sum_{\mathbf{a}} \mathcal{C}^r_{kh}(p;\mathbf{a}), \tag{2.5}$$

where the summation is over $(a_1,\dots,a_t) \in \mathbb{Z}^t$ with

$$a_j \equiv u_j \pmod{p^{h\min\{k_j,k\}}} \quad \text{and} \quad 1 \le a_j \le p^{kh} \quad (1 \le j \le d),$$

and

$$a_j \equiv u_j \pmod{p} \quad \text{and} \quad 1 \le a_j \le p \quad (d+1 \le j \le t).$$

Define $\mathcal{I}=\mathcal{I}(\mathbf{k};k,d)$ as in the preamble to this lemma. Then when $1\le j\le d$ and $j\notin\mathcal{I}$ (if any such j exists), we have $\min\{k_j,k\}=k$. Therefore for each $\mathbf{u}$ the total number of choices for $\mathbf{a}$ is $p^{h\alpha}$, where

$$\alpha=\sum_{j\in\mathcal{I}}(k-k_j). \tag{2.6}$$

Now observe that by assigning the variables $b_1,\dots,b_{r-s}$ arbitrarily in the system (2.4), it follows that when $d\le s\le r$ and $\mathbf{a}\in\mathbb{Z}^t$, one has

$$\mathcal{C}_v^r(p;\mathbf{a})\le p^{(r-s)v}\max_{\mathbf{b}\in\mathbb{Z}^{r-s}}\mathcal{C}_v^s(p;\mathbf{a};\mathbf{b}).$$

Thus by (2.5),

$$\text{card}\left(\mathcal{B}_{k,d}^{\mathbf{k}^*}(p^h;\mathbf{u};\mathbf{f})\right)\le p^{h(\alpha+k(r-d))}\max_{\mathbf{a}\in\mathbb{Z}^t}\max_{\mathbf{b}\in\mathbb{Z}^{r-d}}\mathcal{C}_{kh}^d(p;\mathbf{a};\mathbf{b}). \tag{2.7}$$

Moreover, for each $\mathbf{b}\in\mathbb{Z}^{r-d}$, it follows from Lemma 2.1 that $\mathcal{C}_{kh}^d(p;\mathbf{a};\mathbf{b})\ll_{\mathbf{k}}1$, and thus (2.3) follows from (2.2), (2.6) and (2.7). This completes the proof of the lemma.

Next we provide an estimate related to the number of real singular solutions of a system of additive equations. Before stating our estimate in Lemma 2.3 below, we require some notation. Suppose that t and r are positive integers with $r\le t$, and that $\psi_1,\dots,\psi_t$ are polynomials in $\mathbb{Z}[x]$ satisfying the condition

$$\deg(\psi_1)>\deg(\psi_2)>\dots>\deg(\psi_r)>0.$$

When $\mathcal{I}$ and $\mathcal{J}$ are sets with $\mathcal{I}\subset\{1,2,\dots,2r\}$, $\mathcal{J}\subseteq\{1,2,\dots,t\}$ and $\text{card}(\mathcal{I})=\text{card}(\mathcal{J})$, define the Jacobian determinant $J(\mathcal{I},\mathcal{J};\boldsymbol{\psi})$ by

$$J(\mathcal{I},\mathcal{J};\boldsymbol{\psi})=\det\left(\psi_j'(z_i)\right)_{i\in\mathcal{I},j\in\mathcal{J}}.$$

When d is an integer with $1\le d\le r$, denote by $\mathcal{I}_d^*$ the set of all subsets $\{j_1,\dots,j_d\}$ of $\{1,2,\dots,2r\}$ with $1\le j_1<j_2<\dots<j_d\le 2r$, and write $\mathcal{J}_d$ for the set $\{1,2,\dots,d\}$. We will say that the $2r$-tuple of integers $\mathbf{z}=(z_1,\dots,z_{2r})$ is *highly singular for* $\boldsymbol{\psi}$ if for each $\mathcal{I}\in\mathcal{I}_r^*$ one has $J(\mathcal{I},\mathcal{J}_r;\boldsymbol{\psi})=0$. Further, when $1\le d\le r-1$, we will say that $\mathbf{z}$ is *of type d with respect to* $\boldsymbol{\psi}$ if for some $\mathcal{I}\in\mathcal{I}_d^*$ we have

$$J(\mathcal{I},\mathcal{J}_d;\boldsymbol{\psi})\ne 0, \tag{2.8}$$

and in addition, for each $i\in\{1,2,\dots,2r\}\setminus\mathcal{I}$ we have

$$J(\mathcal{I}\cup\{i\},\mathcal{J}_{d+1};\boldsymbol{\psi})=0. \tag{2.9}$$

We remark that it is an easy exercise, using standard properties of determinants, to show that the type of a $2r$-tuple $\mathbf{z}$ is unique, which is to say that $\mathbf{z}$ cannot be of type d, and of type d', with $d\ne d'$. Finally, we denote by $\mathcal{S}_r(P;\boldsymbol{\psi})$ the set of $2r$-tuples $(z_1,\dots,z_{2r})$, with

$$1\le z_i\le P\quad(1\le i\le 2r), \tag{2.10}$$

which are highly singular for $\boldsymbol{\psi}$.

Lemma 2.3. *Adopt the notation of the previous paragraph, and denote by k_i the degree of ψ_i $(1\le i\le t)$. Then card$(\mathcal{S}_r(P;\boldsymbol{\psi}))\ll_{r,\mathbf{k}}P^{r-1}$.*

Proof. When d is an integer with $1\le d\le r-1$, denote by $\mathcal{T}_d(P;\boldsymbol{\psi})$ the set of $2r$-tuples $\mathbf{z}$, satisfying (2.10), of type d with respect to $\boldsymbol{\psi}$. Further, denote by $\mathcal{T}_0(P;\boldsymbol{\psi})$ the set of $2r$-tuples $\mathbf{z}$ for which $J(\mathcal{I},\mathcal{J}_1;\boldsymbol{\psi})=0$ for each $\mathcal{I}\in\mathcal{I}_1^*$. Consider a $2r$-tuple $\mathbf{z}$ counted by $\mathcal{S}_r(P;\boldsymbol{\psi})$ which does not lie in $\mathcal{T}_0(P;\boldsymbol{\psi})$. Then $J(\mathcal{I},\mathcal{J}_r;\boldsymbol{\psi})=0$ for each $\mathcal{I}\in\mathcal{I}_r^*$, and there is some $\mathcal{I}\in\mathcal{I}_1^*$ with $J(\mathcal{I},\mathcal{J}_1;\boldsymbol{\psi})\ne 0$. It follows from the definition that $\mathbf{z}$ is of type d with respect to $\boldsymbol{\psi}$ for some integer d with $1\le d\le r-1$, and hence that

$$\text{card}\left(\mathcal{S}_r(P;\boldsymbol{\psi})\right)\le\sum_{d=0}^{r-1}\text{card}\left(\mathcal{T}_d(P;\boldsymbol{\psi})\right). \tag{2.11}$$

Observe next that if $\mathbf{z} \in \mathcal{T}_0(P;\boldsymbol{\psi})$, then necessarily

$$\psi_1'(z_i) = 0 \quad (1 \le i \le 2r). \tag{2.12}$$

But $\psi_1(z)$ is a polynomial with degree at least 1, so that $\psi_1'(z)$ is a non-trivial polynomial with at most $k_1 - 1$ roots. It therefore follows from (2.12) that

$$\text{card}\,(\mathcal{T}_0(P;\boldsymbol{\psi})) \le (k_1 - 1)^{2r} \ll_{r,\mathbf{k}} 1. \tag{2.13}$$

Next let d be an integer with $1 \le d \le r-1$, and consider those $\mathbf{z}$ satisfying (2.10) of type d with respect to $\boldsymbol{\psi}$. The number of subsets of $\{1,2,\dots,2r\}$ of cardinality d is $\binom{2r}{d}$. Fix $\mathcal{I}$ to be any one such subset, and suppose that $i \in \{1,2,\dots,2r\} \setminus \mathcal{I}$. A trivial estimate shows that the number of choices of z_l $(l \in \mathcal{I})$ satisfying (2.8) is at most P^d. Fix any one such choice of z_l $(l \in \mathcal{I})$, and suppose that the leading coefficient of the polynomial $\psi_{d+1}'(z)$ is A. Then the leading coefficient, with respect to z_i, of $J(\mathcal{I} \cup \{i\}, \mathcal{J}_{d+1};\boldsymbol{\psi})$ is $AJ(\mathcal{I},\mathcal{J}_d;\boldsymbol{\psi})$. By (2.8), therefore, the equation (2.9) is non-trivial in z_i, and hence the number of possible choices for z_i is at most

$$\deg\left(\psi_{d+1}'(z)\right) = k_{d+1} - 1.$$

Moreover such holds for each $i \in \{1,2,\dots,2r\} \setminus \mathcal{I}$. Consequently, for each d with $1 \le d \le r-1$ one has

$$\text{card}\,(\mathcal{T}_d(P;\boldsymbol{\psi})) \ll_{r,\mathbf{k}} P^d. \tag{2.14}$$

The lemma now follows immediately on combining (2.11), (2.13) and (2.14).

Finally, we recall an estimate for the number of integers in an interval with a given square-free kernel. Given an integer v with canonical prime factorisation $\prod_{i=1}^t p_i^{r_i}$, denote by $s_0(v)$ the square-free kernel of v, that is $\prod_{i=1}^t p_i$.

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

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

Proof. This is Lemma 2.1 of Wooley [23].

3. The fundamental lemma

We aim to establish a fundamental lemma of a form similar to that of [22] (see [23], Lemma 2.2). We first record some notation and conventions. We use vector notation for brevity; for example $(c_1,\dots,c_t)$ will be abbreviated to $\mathbf{c}$. We write $[x]$ for the greatest integer not exceeding x. Also, we use p to denote a prime number, and write $p^s \| n$ when $p^s | n$ but $p^{s+1} \nmid n$. We take $k_1,\dots,k_t$ to be fixed positive integers satisfying (1.2). Throughout, s will denote a positive integer, and ε and η will denote sufficiently small positive numbers. We take P to be a large positive real number depending at most on $\mathbf{k}$, s, ε and η. The implicit constants in Vinogradov's well-known notation, $\ll$ and $\gg$, will depend at most on $\mathbf{k}$, s, ε and η. We write $f \asymp g$ to denote that $f \ll g$ and $g \ll f$. We adopt the following convention concerning the numbers ε and R. Whenever ε or R appear in a statement, either implicitly or explicitly, we assert that for each $\varepsilon > 0$, there exists a positive number $\eta(\varepsilon, s, \mathbf{k})$ such that the statement holds whenever $R \le P^\eta$. Note that the "value" of ε, and of η, may change from statement to statement, and hence also the dependency of implicit constants on ε and η. We observe that since our methods will involve only a finite number of statements (depending at most on $\mathbf{k}$, s and ε), there is no danger of losing control of implicit constants through the successive changes implicit in our arguments.

Let r be a positive integer with $1 \le r \le t$, and let $\Psi_i(z;\mathbf{c})$ $(1 \le i \le t)$ denote polynomials with integer coefficients in the variables $z, c_1, \dots, c_u$. Suppose further that with respect to z, the polynomials $\Psi_i(z;\mathbf{c})$ have non-vanishing leading coefficients, and satisfy the condition

$$\deg\left(\Psi_i(z;\mathbf{c})\right) > \deg\left(\Psi_{i+1}(z;\mathbf{c})\right) \quad (1 \le i < r).$$

Define the Jacobian $J_r(\mathbf{z};\mathbf{c})$ by

$$J_r(\mathbf{z};\mathbf{c}) = \det\left(\frac{\partial \Psi_i}{\partial z_j}(z_j;\mathbf{c})\right)_{1 \le i,j \le r}.$$

Let Q be a real number with $R \le Q \le P$, and let C_i, C_i' $(1 \le i \le u)$ be real numbers with $1 \le C_i' \le C_i \ll P$. We write

$$\widetilde{C} = \prod_{i=1}^{u} C_i.$$

Let $D_1(\mathbf{c}), \dots, D_t(\mathbf{c})$ denote polynomials in $\mathbb{Z}[c_1, \dots, c_u]$ with total degrees bounded in terms of $\mathbf{k}$, and satisfying the property that $D_i(\mathbf{c}) \ne 0$ for $C_i' < c_i \le C_i$ $(1 \le i \le u)$. Denote by

$$S_{s,r}(P, Q, R) = S_{s,r}^{(\mathbf{k})}(P, Q, R; \mathbf{\Psi}; \mathbf{C}, \mathbf{C}'; \mathbf{D}) \tag{3.1}$$

the number of solutions of the simultaneous equations

$$\sum_{n=1}^{r} \eta_n \left(\Psi_i(z_n; \mathbf{c}) - \Psi_i(w_n; \mathbf{c})\right) + D_i(\mathbf{c}) \sum_{m=1}^{s} (x_m^{k_i} - y_m^{k_i}) = 0 \quad (1 \le i \le t), \tag{3.2}$$

with

$$x_m, y_m \in \mathcal{A}(Q, R) \quad (1 \le m \le s), \quad C_j' < c_j \le C_j \quad (1 \le j \le u), \tag{3.3}$$

$$1 \le z_n, w_n \le P \quad \text{and} \quad \eta_n \in \{+1, -1\} \quad (1 \le n \le r). \tag{3.4}$$

Further, denote by

$$\widetilde{S}_{s,r}(P, Q, R) = \widetilde{S}_{s,r}^{(\mathbf{k})}(P, Q, R; \mathbf{\Psi}; \mathbf{C}, \mathbf{C}'; \mathbf{D})$$

the number of solutions of the system (3.2) with (3.3), (3.4), and the additional conditions $J_r(\mathbf{z}; \mathbf{c}) \ne 0$ and $J_r(\mathbf{w}; \mathbf{c}) \ne 0$. For a given real number θ with $1 < P^\theta < Q^{1/2-\varepsilon}$, let

$$T_{s,r}(P, Q, R; \theta) = T_{s,r}^{(\mathbf{k})}(P, Q, R; \theta; \mathbf{\Psi}; \mathbf{C}, \mathbf{C}'; \mathbf{D}) \tag{3.5}$$

denote the number of solutions of the simultaneous diophantine equations

$$\sum_{n=1}^{r} \eta_n \left(\Psi_i(z_n; \mathbf{c}) - \Psi_i(w_n; \mathbf{c})\right) + D_i(\mathbf{c}) q^{k_i} \sum_{m=1}^{s} (u_m^{k_i} - v_m^{k_i}) = 0 \quad (1 \le i \le t), \tag{3.6}$$

with $\mathbf{z}, \mathbf{w}, \mathbf{c}, \boldsymbol{\eta}$ satisfying (3.3), (3.4), and

$$P^\theta < q \le P^\theta R, \quad u_m, v_m \in \mathcal{A}(QP^{-\theta}, R) \quad (1 \le m \le s), \tag{3.7}$$

and

$$(J_r(\mathbf{z}; \mathbf{c}), q) = (J_r(\mathbf{w}; \mathbf{c}), q) = 1. \tag{3.8}$$

In Lemma 3.1 below, through a substantial elaboration of the argument of the proof of [23], Lemma 2.2, we bound $S_{s,r}$ in terms of $T_{s,r}$ and $\widetilde{S}_{s-1,r}$.

Lemma 3.1. *Let $s \in \mathbb{N}$, and suppose that $\theta = \theta(r, s, \mathbf{k}; \mathbf{\Psi})$ satisfies $1 < P^\theta < Q^{1/2-\varepsilon}$. Then there exists a positive number $\eta = \eta(\varepsilon, s, \mathbf{k})$ such that whenever* $\exp((\log \log P)^2) < R \le P^\eta$*,*

$$S_{s,r}(P, Q, R) \ll_{\mathbf{\Psi}} QP^{\theta+\varepsilon} \widetilde{S}_{s-1,r}(P, Q, R) + P^{(2s-1)\theta+\varepsilon} T_{s,r}(P, Q, R; \theta).$$

Proof. In order to facilitate our analysis, we initially classify the solutions of (3.2) counted by $S_{s,r}(P, Q, R)$ into three types. We let S_1 denote the number of solutions of the system (3.2) satisfying (3.3) and (3.4) for which there is a j with

$$\min\{x_j, y_j\} \le P^\theta, \tag{3.9}$$

let S_2 denote the number of solutions for which $\mathbf{z}, \mathbf{w}$ is highly singular for $\mathbf{\Psi}$, and for which (3.9) holds for no j, and let S_3 denote the number of solutions for which $\mathbf{z}, \mathbf{w}$ is not highly singular for $\mathbf{\Psi}$, and for which (3.9) holds for no j. Then plainly

$$S_{s,r}(P, Q, R) \le 3 \max\{S_1, S_2, S_3\}.$$

We divide into cases.

(i) Suppose that $S_1 = \max\{S_1, S_2, S_3\}$, so that $S_{s,r}(P,Q,R) \le 3S_1$. Write

$$\psi(\boldsymbol{\alpha}; x) = \sum_{i=1}^{t} \alpha_i D_i(\mathbf{c}) x^{k_i} \quad \text{and} \quad \Phi(\boldsymbol{\alpha}; z, \mathbf{c}) = \sum_{i=1}^{t} \alpha_i \Psi_i(z; \mathbf{c}).$$

Next define the exponential sums $f_{\mathbf{c}}$ and $F_{\mathbf{c}}$ by

$$f_{\mathbf{c}}(\boldsymbol{\alpha}; L, R) = \sum_{x \in \mathcal{A}(L,R)} e(\psi(\boldsymbol{\alpha}; x)) \quad \text{and} \quad F_{\mathbf{c}}(\boldsymbol{\alpha}; P, R) = \sum_{1 \le z \le P} e(\Phi(\boldsymbol{\alpha}; z, \mathbf{c})),$$

and write

$$\widetilde{F}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha}; P, R) = \prod_{n=1}^{r} F_{\mathbf{c}}(\eta_n \boldsymbol{\alpha}; P, R).$$

Then by considering the underlying diophantine equations, we obtain

$$S_1 \ll \sum_{\mathbf{c},\boldsymbol{\eta}} \int_{\mathbb{T}^t} \left| \widetilde{F}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha}; P, R)^2 f_{\mathbf{c}}(\boldsymbol{\alpha}; P^\theta, R) f_{\mathbf{c}}(\boldsymbol{\alpha}; Q, R)^{2s-1} \right| d\boldsymbol{\alpha},$$

where the summation is over $\boldsymbol{\eta}$ satisfying (3.4) and $\mathbf{c}$ satisfying (3.3). Write $\mathbf{1}$ for the vector $\boldsymbol{\eta} = (1, \dots, 1)$. Then by Hölder's inequality,

$$S_{s,r}(P,Q,R) \ll I_1(P^\theta)^{\frac{1}{2s}} I_1(Q)^{1-\frac{1}{2s}},$$

where

$$I_1(L) = \sum_{\mathbf{c}} \int_{\mathbb{T}^t} \left| \widetilde{F}_{\mathbf{c},\mathbf{1}}(\boldsymbol{\alpha}; P, R)^2 f_{\mathbf{c}}(\boldsymbol{\alpha}; L, R)^{2s} \right| d\boldsymbol{\alpha}.$$

By considering the underlying diophantine equations, we therefore deduce that

$$S_{s,r}(P,Q,R) \ll \left(S_{s,r}(P,P^\theta,R)\right)^{\frac{1}{2s}} \left(S_{s,r}(P,Q,R)\right)^{1-\frac{1}{2s}},$$

so that $S_{s,r}(P,Q,R) \ll S_{s,r}(P,P^\theta,R)$. Now observe that by counting only those solutions of (3.2) for which $x_m = y_m$ $(1 \le m \le s)$, we obtain

$$S_{s,r}(P,Q,R) \gg Q^{s-\varepsilon} S_{0,r}(P,Q,R).$$

On the other hand, counting the possible choices for x_m, y_m $(1 \le m \le s)$ trivially, and considering the mean value corresponding to (3.2), one obtains

$$S_{s,r}(P,P^\theta,R) \ll P^{2s\theta} S_{0,r}(P,P^\theta,R) = P^{2s\theta} S_{0,r}(P,Q,R).$$

Thus, since $P^\theta < Q^{1/2-\varepsilon}$, we obtain

$$Q^{s-\varepsilon} S_{0,r}(P,Q,R) \ll S_{s,r}(P,Q,R) \ll Q^{s-2s\varepsilon} S_{0,r}(P,Q,R),$$

which provides a contradiction, since P and Q are sufficiently large. Thus we may suppose that $S_1 \ll \max\{S_2, S_3\}$.

(ii) Suppose that $S_2 = \max\{S_2, S_3\}$. Then the conclusion of the preceding paragraph implies that $S_{s,r}(P,Q,R) \ll S_2$. By Lemma 2.3, the number of $\mathbf{z}$ and $\mathbf{w}$ satisfying (3.4), for which $\mathbf{z}, \mathbf{w}$ is highly singular for $\boldsymbol{\Psi}$, is $O(P^{r-1})$. Fix any one such choice of $\mathbf{z}$, $\mathbf{w}$, and any choice of $\mathbf{c}$ and $\boldsymbol{\eta}$. Then by considering the mean value corresponding to (3.2), and considering the underlying diophantine equations, one deduces that the number of choices for $\mathbf{x}$ and $\mathbf{y}$, as one ranges across solutions $\mathbf{z}, \mathbf{w}, \mathbf{c}, \mathbf{x}, \mathbf{y}, \boldsymbol{\eta}$ counted by S_2, is bounded above by

$$\int_{\mathbb{T}^t} |f_{\mathbf{c}}(\boldsymbol{\alpha}; Q, R)|^{2s} d\boldsymbol{\alpha} = S_s(Q,R).$$

It follows that
$$S_2 \ll P^{r-1}\widetilde{C}S_s(Q,R).$$
However, by counting only those solutions in which $z_n = w_n$ $(1 \le n \le r)$, we find that
$$S_{s,r}(P,Q,R) \gg P^r\widetilde{C}S_s(Q,R).$$
Consequently,
$$P^r\widetilde{C}S_s(Q,R) \ll S_{s,r}(P,Q,R) \ll S_2 \ll P^{r-1}\widetilde{C}S_s(Q,R),$$
which again leads to a contradiction, since P is sufficiently large. Thus we may suppose that $S_2 \ll S_3$.

(iii) We now estimate S_3. In view of the conclusions of the two previous cases, we may suppose that $S_{s,r}(P,Q,R) \ll S_3$. Consider any solution $\mathbf{z},\mathbf{w},\mathbf{c},\mathbf{x},\mathbf{y},\boldsymbol{\eta}$ with $\mathbf{z},\mathbf{w}$ not highly singular for $\boldsymbol{\Psi}$, and for which (3.9) holds for no j. Define the $2r$-tuple $\boldsymbol{\zeta}$ by writing $\zeta_i = z_i$ and $\zeta_{r+i} = w_i$ for $1 \le i \le r$. By the definition of highly singular, it follows that there exists a subset $\mathcal{G}$ of $\{1,2,\dots,2r\}$ with $\mathrm{card}(\mathcal{G}) = r$, and satisfying the property that
$$\det\left(\Psi_j'(\zeta_i;\mathbf{c})\right)_{1\le j\le r,\ i\in\mathcal{G}} \ne 0.$$
Thus, by a rearrangement of variables, it follows that
$$S_{s,r}(P,Q,R) \ll S_4, \tag{3.10}$$
where S_4 denotes the number of solutions of (3.2) with (3.3) and (3.4) in which $J_r(\mathbf{z};\mathbf{c}) \ne 0$, and in which (3.9) holds for no j. Notice that we do not preclude the possibility that $J_r(\mathbf{w};\mathbf{c}) = 0$.

Write
$$\widetilde{F}^*_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha};P,R) = \sum_{\mathbf{z}} e(\eta_1\Phi(\boldsymbol{\alpha};z_1,\mathbf{c}) + \cdots + \eta_r\Phi(\boldsymbol{\alpha};z_r,\mathbf{c})),$$
where the summation is over $\mathbf{z}$ satisfying (3.4), with the additional condition that $J_r(\mathbf{z};\mathbf{c}) \ne 0$. Also, write
$$f^*_{\mathbf{c}}(\boldsymbol{\alpha};Q,R;L) = \sum_{\substack{x\in\mathcal{A}(Q,R)\\ x>L}} e(\psi(\boldsymbol{\alpha};x)).$$
Then by considering the underlying diophantine equations, it follows from (3.10) that
$$S_{s,r}(P,Q,R) \ll \sum_{\mathbf{c},\boldsymbol{\eta},\boldsymbol{\omega}} \int_{\mathbb{T}^t} \left|\widetilde{F}^*_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha};P,R)\widetilde{F}_{\mathbf{c},\boldsymbol{\omega}}(\boldsymbol{\alpha};P,R)f^*_{\mathbf{c}}(\boldsymbol{\alpha};Q,R;P^\theta)^{2s}\right| d\boldsymbol{\alpha},$$
where the summation is over $\boldsymbol{\eta},\boldsymbol{\omega} \in \{+1,-1\}^r$, and $\mathbf{c}$ satisfying (3.3). Thus, by the Cauchy-Schwarz inequalities,
$$S_{s,r}(P,Q,R) \ll I_1^{1/2}I_2^{1/2}, \tag{3.11}$$
where
$$I_1 = \sum_{\mathbf{c},\boldsymbol{\eta}} \int_{\mathbb{T}^t} \left|\widetilde{F}^*_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha};P,R)^2 f^*_{\mathbf{c}}(\boldsymbol{\alpha};Q,R;P^\theta)^{2s}\right| d\boldsymbol{\alpha},$$
and
$$I_2 = \sum_{\mathbf{c},\boldsymbol{\omega}} \int_{\mathbb{T}^t} \left|\widetilde{F}_{\mathbf{c},\boldsymbol{\omega}}(\boldsymbol{\alpha};P,R)^2 f^*_{\mathbf{c}}(\boldsymbol{\alpha};Q,R;P^\theta)^{2s}\right| d\boldsymbol{\alpha}.$$
Moreover, by considering the underlying diophantine equations, one has
$$I_2 \ll S_{s,r}(P,Q,R).$$
On substituting the latter bound into (3.11) we deduce that
$$S_{s,r}(P,Q,R) \ll I_1. \tag{3.12}$$
Also, on considering the underlying diophantine equations, we have
$$I_1 \le I_3, \tag{3.13}$$

where I_3 denotes the number of solutions of (3.2) with (3.3) and (3.4) in which $J_r(\mathbf{z};\mathbf{c})\neq 0$, $J_r(\mathbf{w};\mathbf{c})\neq 0$, and in which (3.9) holds for no j.

We now classify the solutions counted by I_3 into two types. For the sake of convenience, we write $x\mathcal{D}(L)y$ to denote that there is some divisor d of x with $d\le L$ such that x/d has all of its prime divisors amongst those of y. Let S_5 denote the number of solutions counted by I_3 for which

$$x_j\mathcal{D}(P^\theta)J_r(\mathbf{z};\mathbf{c})\quad \text{or}\quad y_j\mathcal{D}(P^\theta)J_r(\mathbf{w};\mathbf{c}) \tag{3.14}$$

for at least one j, and let S_6 denote the number of solutions counted by I_3 for which (3.14) holds for no j. Then it follows from (3.12) and (3.13) that

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

We divide into further cases.

(iv) Suppose that $S_5=\max\{S_5,S_6\}$, so that $S_{s,r}(P,Q,R)\ll S_5$. Given $\mathbf{z}$ and $\mathbf{c}$ satisfying (3.3) and (3.4) with $J_r(\mathbf{z};\mathbf{c})\neq 0$, denote by $\mathcal{S}(\mathbf{z};\mathbf{c})$ the set of positive integers x such that $x\le Q$, and x has a divisor d with $d\le P^\theta$ satisfying the condition that x/d has all of its prime divisors amongst those of $J_r(\mathbf{z};\mathbf{c})$. Let

$$\widetilde{H}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha};P,Q,R)=\sum_{\mathbf{z}}\sum_{x\in\mathcal{S}(\mathbf{z};\mathbf{c})}e(\Xi(\boldsymbol{\alpha};x,\mathbf{z},\mathbf{c};\boldsymbol{\eta})),$$

where the summation is over $\mathbf{z}$ satisfying (3.4) with $J_r(\mathbf{z};\mathbf{c})\neq 0$, and

$$\Xi(\boldsymbol{\alpha};x,\mathbf{z},\mathbf{c};\boldsymbol{\eta})=\sum_{i=1}^t\alpha_i\left(D_i(\mathbf{c})x^{k_i}+\eta_1\Psi_i(z_1;\mathbf{c})+\cdots+\eta_r\Psi_i(z_r;\mathbf{c})\right).$$

Then

$$S_5\ll\sum_{\mathbf{c},\boldsymbol{\eta},\boldsymbol{\omega}}\int_{\mathbb{T}^t}\left|\widetilde{H}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha};P,Q,R)\widetilde{F}^*_{\mathbf{c},\boldsymbol{\omega}}(\boldsymbol{\alpha};P,R)f^*_{\mathbf{c}}(\boldsymbol{\alpha};Q,R;P^\theta)^{2s-1}\right|d\boldsymbol{\alpha},$$

so that by Schwarz's inequality, on considering the underlying diophantine equations,

$$S_5\ll(S_{s,r}(P,Q,R))^{1/2}\,I_4^{1/2}, \tag{3.15}$$

where

$$I_4=\sum_{\mathbf{c},\boldsymbol{\eta}}\int_{\mathbb{T}^t}\left|\widetilde{H}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha};P,Q,R)^2f^*_{\mathbf{c}}(\boldsymbol{\alpha};Q,R;P^\theta)^{2s-2}\right|d\boldsymbol{\alpha}.$$

Thus, by considering the underlying diophantine equations, we deduce from (3.15) that

$$S_{s,r}(P,Q,R)\ll\sum_{g,h}\sum_{\mathbf{c}}V(g,h;\mathbf{c}), \tag{3.16}$$

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

$$\begin{aligned}\sum_{n=1}^r\eta_n\left(\Psi_i(z_n;\mathbf{c})-\Psi_i(w_n;\mathbf{c})\right)+D_i(\mathbf{c})&\sum_{m=1}^{s-1}\left(x_m^{k_i}-y_m^{k_i}\right)\\ &=D_i(\mathbf{c})\left((ey)^{k_i}-(dx)^{k_i}\right)\quad(1\le i\le t),\end{aligned}$$

with $\mathbf{z},\mathbf{w},\mathbf{x},\mathbf{y},\boldsymbol{\eta},d,e$ satisfying (3.3) and (3.4), and subject to

$$J_r(\mathbf{z};\mathbf{c})\neq 0,\quad J_r(\mathbf{w};\mathbf{c})\neq 0,\quad g|J_r(\mathbf{z};\mathbf{c}),\quad h|J_r(\mathbf{w};\mathbf{c}),$$

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

Let

$$G_{\mathbf{c},\boldsymbol{\eta},g}(\boldsymbol{\alpha};P,R)=\sum_{\mathbf{z}}e(\Xi(\boldsymbol{\alpha};0,\mathbf{z},\mathbf{c};\boldsymbol{\eta})),$$

where the summation is over $\mathbf{z}$ satisfying (3.4), and subject to $J_r(\mathbf{z};\mathbf{c})\ne 0$ and $g|J_r(\mathbf{z};\mathbf{c})$. Write

$$\mathcal{G}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha})=\sum_{g\le Q}G_{\mathbf{c},\boldsymbol{\eta},g}(\boldsymbol{\alpha};P,R)\sum_{d\le P^\theta}\sum_{\substack{x\le Q/d\\ s_0(x)=g}}e\left(\sum_{i=1}^t\alpha_iD_i(\mathbf{c})(dx)^{k_i}\right). \tag{3.17}$$

Here, if g is not squarefree, we understand the third summation in (3.17) to be empty. Then by (3.16), on considering the underlying diophantine equations,

$$S_{s,r}(P,Q,R)\ll\sum_{\mathbf{c},\boldsymbol{\eta}}\int_{\mathbb{T}^t}\left|\mathcal{G}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha})^2f_{\mathbf{c}}(\boldsymbol{\alpha};Q,R)^{2s-2}\right|d\boldsymbol{\alpha}. \tag{3.18}$$

But by Cauchy's inequality,

$$|\mathcal{G}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha})|^2\le\mathcal{H}_{1,\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha})\mathcal{H}_{2,\mathbf{c}}(\boldsymbol{\alpha}), \tag{3.19}$$

where

$$\mathcal{H}_{1,\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha})=\sum_{g\le Q}|G_{\mathbf{c},\boldsymbol{\eta},g}(\boldsymbol{\alpha};P,R)|^2, \tag{3.20}$$

and

$$\mathcal{H}_{2,\mathbf{c}}(\boldsymbol{\alpha})=\sum_{g\le Q}\left|\sum_{d\le P^\theta}\sum_{\substack{x\le Q/d\\ s_0(x)=g}}e\left(\sum_{i=1}^t\alpha_iD_i(\mathbf{c})(dx)^{k_i}\right)\right|^2.$$

Moreover, by interchanging the order of summation, and applying Cauchy's inequality in combination with Lemma 2.4 (as in the argument of the proof of [23], Lemma 2.2 part (iii)), we obtain

$$\begin{aligned}\mathcal{H}_{2,\mathbf{c}}(\boldsymbol{\alpha})&=\sum_{g\le Q}\left|\sum_{\substack{x\le Q\\ s_0(x)=g}}\sum_{\substack{d\le P^\theta\\ d\le Q/x}}e\left(\sum_{i=1}^t\alpha_iD_i(\mathbf{c})(dx)^{k_i}\right)\right|^2\\ &\ll P^\varepsilon\sum_{g\le Q}\sum_{\substack{x\le Q\\ s_0(x)=g}}P^\theta Q/x\\ &\ll QP^{\theta+\varepsilon}.\end{aligned} \tag{3.21}$$

On combining (3.18)-(3.21), we deduce that

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

But by considering the underlying diophantine equation, and using standard estimates for the divisor function, we find that the integral on the right hand side of the last inequality is

$$\ll P^\varepsilon\widetilde{S}_{s-1,r}(P,Q,R).$$

The desired conclusion now follows in the fourth case.

(v) Suppose that $S_6\ge\max\{S_5,S_6\}$, so that $S_{s,r}(P,Q,R)\ll S_6$. For a given solution of (3.2) satisfying (3.3) and (3.4) counted by S_6, we have for every j,

$$x_j>P^\theta\quad\text{and}\quad y_j>P^\theta, \tag{3.22}$$

$$J_r(\mathbf{z};\mathbf{c})\ne0,\quad J_r(\mathbf{w};\mathbf{c})\ne0, \tag{3.23}$$

and neither

$$x_j\mathcal{D}(P^\theta)J_r(\mathbf{z};\mathbf{c})\quad\text{nor}\quad y_j\mathcal{D}(P^\theta)J_r(\mathbf{w};\mathbf{c}). \tag{3.24}$$

Consider a fixed index j. Let q be the greatest divisor of x_j with the property that $(q, J_r(\mathbf{z};\mathbf{c})) = 1$. If $q \le P^\theta$, then $x_j \mathcal{D}(P^\theta) J_r(\mathbf{z};\mathbf{c})$, contradicting (3.24). Hence $q > P^\theta$, and since each prime divisor of x_j is at most R, there exists a divisor q_j of x_j with $P^\theta < q_j \le P^\theta R$, and satisfying $(q_j, J_r(\mathbf{z};\mathbf{c})) = 1$. We may do likewise with the y_j. We therefore deduce that $S_6 \ll V_1$, where V_1 denotes the number of solutions of the system of equations

$$\sum_{n=1}^{r} \eta_n \left(\Psi_i(z_n;\mathbf{c}) - \Psi_i(w_n;\mathbf{c})\right) + D_i(\mathbf{c}) \sum_{j=1}^{s} \left((q_j u_j)^{k_i} - (p_j v_j)^{k_i}\right) = 0 \quad (1 \le i \le t),$$

with $\mathbf{z}, \mathbf{w}, \mathbf{c}, \boldsymbol{\eta}$ satisfying (3.3) and (3.4), and for $j = 1, \dots, s$ with

$$P^\theta < q_j, p_j \le P^\theta R, \quad u_j \in \mathcal{A}(Q/q_j, R), \quad v_j \in \mathcal{A}(Q/p_j, R) \tag{3.25}$$

and

$$(q_j, J_r(\mathbf{z};\mathbf{c})) = (p_j, J_r(\mathbf{w};\mathbf{c})) = 1.$$

Let

$$F_{\mathbf{c},\boldsymbol{\eta},q}(\boldsymbol{\alpha}; P, R) = \sum_{\mathbf{z}} e\left(\Xi(\boldsymbol{\alpha}; 0, \mathbf{z}, \mathbf{c}; \boldsymbol{\eta})\right),$$

where the summation is over $\mathbf{z}$ satisfying (3.4) subject to $(q, J_r(\mathbf{z};\mathbf{c})) = 1$, and let

$$\mathcal{F}_{\mathbf{c},j}(\boldsymbol{\alpha}) = f_{\mathbf{c}}(\mathbf{q}_j^k \boldsymbol{\alpha}; Q/q_j, R) f_{\mathbf{c}}(-\mathbf{p}_j^k \boldsymbol{\alpha}; Q/p_j, R),$$

where $\mathbf{q}_j^k \boldsymbol{\alpha}$ denotes $(q_j^{k_1}\alpha_1, \dots, q_j^{k_t}\alpha_t)$, and similarly for $\mathbf{p}_j^k \boldsymbol{\alpha}$. Then by considering the underlying diophantine equations,

$$V_1 \le \sum_{\mathbf{c},\boldsymbol{\eta}} \int_{\mathbb{T}^t} \sum_{\mathbf{q},\mathbf{p}} F_{\mathbf{c},\boldsymbol{\eta},\widetilde{q}}(\boldsymbol{\alpha}; P, R) F_{\mathbf{c},\boldsymbol{\eta},\widetilde{p}}(-\boldsymbol{\alpha}; P, R) \prod_{j=1}^{s} \mathcal{F}_{\mathbf{c},j}(\boldsymbol{\alpha}) d\boldsymbol{\alpha}, \tag{3.26}$$

where the summation is over $\mathbf{q}, \mathbf{p}$ satisfying (3.25) and we have written $\widetilde{q} = q_1 \dots q_s$, and likewise $\widetilde{p} = p_1 \dots p_s$.

Let

$$X_{\mathbf{c},\boldsymbol{\eta},j}(\boldsymbol{\alpha}) = \left| F_{\mathbf{c},\boldsymbol{\eta},\widetilde{q}}(\boldsymbol{\alpha}; P, R)^2 f_{\mathbf{c}}(\mathbf{q}_j^k \boldsymbol{\alpha}; Q/q_j, R)^{2s} \right|$$

and let $Y_{\mathbf{c},\boldsymbol{\eta},j}(\boldsymbol{\alpha})$ denote the analogous function appropriate to the p_j. Then it follows from (3.26) that

$$S_6 \ll \sum_{\mathbf{q},\mathbf{p}} \sum_{\mathbf{c},\boldsymbol{\eta}} \int_{\mathbb{T}^t} \prod_{j=1}^{s} \left(X_{\mathbf{c},\boldsymbol{\eta},j}(\boldsymbol{\alpha}) Y_{\mathbf{c},\boldsymbol{\eta},j}(\boldsymbol{\alpha})\right)^{\frac{1}{2s}} d\boldsymbol{\alpha}. \tag{3.27}$$

By Hölder's inequality,

$$\begin{aligned} &\sum_{\mathbf{c},\boldsymbol{\eta}} \int_{\mathbb{T}^t} \prod_{j=1}^{s} \left(X_{\mathbf{c},\boldsymbol{\eta},j}(\boldsymbol{\alpha}) Y_{\mathbf{c},\boldsymbol{\eta},j}(\boldsymbol{\alpha})\right)^{\frac{1}{2s}} d\boldsymbol{\alpha} \\ &\qquad \ll \prod_{j=1}^{s} \left(\sum_{\mathbf{c},\boldsymbol{\eta}} \int_{\mathbb{T}^t} X_{\mathbf{c},\boldsymbol{\eta},j}(\boldsymbol{\alpha}) d\boldsymbol{\alpha} \right)^{\frac{1}{2s}} \left(\sum_{\mathbf{c},\boldsymbol{\eta}} \int_{\mathbb{T}^t} Y_{\mathbf{c},\boldsymbol{\eta},j}(\boldsymbol{\alpha}) d\boldsymbol{\alpha} \right)^{\frac{1}{2s}}. \end{aligned}$$

Now observe that

$$\sum_{\mathbf{c},\boldsymbol{\eta}} \int_{\mathbb{T}^t} X_{\mathbf{c},\boldsymbol{\eta},j}(\boldsymbol{\alpha}) d\boldsymbol{\alpha} \le W(P, Q, R, q_j),$$

where $W(P, Q, R, q)$ denotes the number of solutions of the system (3.6) with $\mathbf{z}, \mathbf{w}, \mathbf{c}, \boldsymbol{\eta}, \mathbf{u}, \mathbf{v}$ satisfying (3.3), (3.4), (3.7) and (3.8). Therefore by Hölder's inequality,

$$\begin{aligned} &\sum_{\mathbf{q},\mathbf{p}} \sum_{\mathbf{c},\boldsymbol{\eta}} \int_{\mathbb{T}^t} \prod_{j=1}^{s} \left(X_{\mathbf{c},\boldsymbol{\eta},j}(\boldsymbol{\alpha}) Y_{\mathbf{c},\boldsymbol{\eta},j}(\boldsymbol{\alpha})\right)^{\frac{1}{2s}} d\boldsymbol{\alpha} \\ &\qquad \ll \sum_{\mathbf{q},\mathbf{p}} \prod_{j=1}^{s} \left(W(P, Q, R, q_j) W(P, Q, R, p_j)\right)^{\frac{1}{2s}} \\ &\qquad \ll \left(\sum_{\mathbf{q},\mathbf{p}} 1 \right)^{1-\frac{1}{2s}} \left(\sum_{\mathbf{q},\mathbf{p}} \prod_{j=1}^{s} \left(W(P, Q, R, q_j) W(P, Q, R, p_j)\right) \right)^{\frac{1}{2s}} \\ &\qquad \ll \prod_{j=1}^{2s} \left(\left(P^\theta R\right)^{2s-1} T_{s,r}(P, Q, R; \theta) \right)^{\frac{1}{2s}}, \end{aligned} \tag{3.28}$$

where $T_{s,r}(P,Q,R;\theta)$ is defined in the preamble to this lemma. On combining (3.27) and (3.28), we complete the proof of the lemma.

Our next step will be to convert the congruence conditions implicit in the system of equations (3.6) into a congruence condition which is more readily exploited. For a given real number θ with $1 < P^\theta < Q^{1/2-\varepsilon}$, let

$$\widetilde{T}_{s,r,k}(P,Q,R;\theta) = \widetilde{T}^{(\mathbf{k})}_{s,r,k}(P,Q,R;\theta;\mathbf{\Psi};\mathbf{C},\mathbf{C}';\mathbf{D})$$

denote the number of solutions of the system of equations (3.6) with (3.3), (3.4) and (3.7), satisfying the additional condition

$$z_n \equiv w_n \pmod{q^k} \quad (1 \le n \le r).$$

In Lemma 3.2 below we show that Lemma 2.2 can be employed to relate $T_{s,r}$ to $\widetilde{T}_{s,r,k}$ in a simple manner.

Lemma 3.2. *Let $\theta = \theta(r,s,\mathbf{k};\mathbf{\Psi})$ satisfy $1 < P^\theta < Q^{1/2-\varepsilon}$, and suppose that k is a positive integer. Then*

$$T_{s,r}(P,Q,R;\theta) \ll P^{\omega(\mathbf{k};k,r,r)\theta+\varepsilon}\widetilde{T}_{s,r,k}(P,Q,R;\theta).$$

Proof. For a given q satisfying (3.7), let $\mathcal{B}(\mathbf{u};\mathbf{c},\boldsymbol{\eta})$ denote the set of solutions, $\mathbf{z}$, distinct modulo q^k of the system of congruences

$$\Upsilon_i(\mathbf{z};\mathbf{c},\boldsymbol{\eta}) \equiv u_i \pmod{q^{k_i}} \quad (1 \le i \le t),$$

with $(J_r(\mathbf{z};\mathbf{c}),q) = 1$, where

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

Then by Lemma 2.2,

$$\mathrm{card}\,(\mathcal{B}(\mathbf{u};\mathbf{c},\boldsymbol{\eta})) \ll q^{\omega(\mathbf{k};k,r,r)+\varepsilon}, \tag{3.29}$$

where $\omega(\mathbf{k};k,r,r)$ is defined by (2.2). Moreover for each solution $\mathbf{z},\mathbf{w},\mathbf{c},\boldsymbol{\eta},q,\mathbf{u},\mathbf{v}$ counted by $T_{s,r}(P,Q,R;\theta)$, it follows from (3.6) that

$$\Upsilon_i(\mathbf{z};\mathbf{c},\boldsymbol{\eta}) \equiv \Upsilon_i(\mathbf{w};\mathbf{c},\boldsymbol{\eta}) \pmod{q^{k_i}} \quad (1 \le i \le t).$$

Thus each solution of (3.6) may be classified according to the common residue class modulo q^{k_i} of $\Upsilon_i(\mathbf{z};\mathbf{c},\boldsymbol{\eta})$ and $\Upsilon_i(\mathbf{w};\mathbf{c},\boldsymbol{\eta})$ for $1 \le i \le t$.

Let

$$H(\boldsymbol{\alpha};\mathbf{z};\mathbf{c},\boldsymbol{\eta}) = \sum_{\substack{1\le x_1\le P\\ x_1\equiv z_1 \pmod{q^k}}} \cdots \sum_{\substack{1\le x_r\le P\\ x_r\equiv z_r \pmod{q^k}}} e\left(\sum_{i=1}^t \alpha_i \Upsilon_i(\mathbf{x};\mathbf{c},\boldsymbol{\eta})\right).$$

Then we have

$$T_{s,r}(P,Q,R;\theta) \ll \sum_q \sum_{\mathbf{c},\boldsymbol{\eta}} \int_{\mathbb{T}^t} \widetilde{H}(\boldsymbol{\alpha};\mathbf{c},\boldsymbol{\eta}) \left|\widetilde{f}_{\mathbf{c},q}(\boldsymbol{\alpha};QP^{-\theta},R)\right|^{2s} d\boldsymbol{\alpha}, \tag{3.30}$$

where the summation is over $\mathbf{c},\boldsymbol{\eta}$ satisfying (3.3), (3.4), and q satisfying (3.7), and we have written

$$\widetilde{H}(\boldsymbol{\alpha};\mathbf{c},\boldsymbol{\eta}) = \sum_{u_1=1}^{q^{k_1}} \cdots \sum_{u_t=1}^{q^{k_t}} \left|\sum_{\mathbf{z}\in\mathcal{B}(\mathbf{u};\mathbf{c},\boldsymbol{\eta})} H(\boldsymbol{\alpha};\mathbf{z};\mathbf{c},\boldsymbol{\eta})\right|^2,$$

$$\widetilde{f}_{\mathbf{c},q}(\boldsymbol{\alpha};L,R) = \sum_{x\in\mathcal{A}(L,R)} e\left(\sum_{i=1}^t \alpha_i D_i(\mathbf{c})(qx)^{k_i}\right).$$

But by Cauchy's inequality,

$$\widetilde{H}(\boldsymbol{\alpha};\mathbf{c},\boldsymbol{\eta}) \le \sum_{u_1=1}^{q^{k_1}} \cdots \sum_{u_t=1}^{q^{k_t}} \mathrm{card}\,(\mathcal{B}(\mathbf{u};\mathbf{c},\boldsymbol{\eta})) \sum_{\mathbf{z}\in\mathcal{B}(\mathbf{u};\mathbf{c},\boldsymbol{\eta})} |H(\boldsymbol{\alpha};\mathbf{z};\mathbf{c},\boldsymbol{\eta})|^2.$$

Then by (3.29) and (3.30),

$$T_{s,r}(P,Q,R;\theta) \ll (P^\theta R)^{\omega(\mathbf{k};k,r,r)+\varepsilon} \sum_q \sum_{\mathbf{c},\boldsymbol{\eta}} \mathcal{T}(q,\mathbf{c},\boldsymbol{\eta}),$$

where

$$\mathcal{T}(q,\mathbf{c},\boldsymbol{\eta}) = \sum_{z_1=1}^{q^k} \cdots \sum_{z_r=1}^{q^k} \int_{\mathbb{T}^t} \left| H(\boldsymbol{\alpha};\mathbf{z};\mathbf{c},\boldsymbol{\eta})^2 \widetilde{f}_{\mathbf{c},q}(\boldsymbol{\alpha};QP^{-\theta},R)^{2s} \right| d\boldsymbol{\alpha}.$$

The lemma now follows on considering the underlying diophantine equations.

4. Efficient differencing

We now develop the efficient differencing process which is fundamental to the new mean value estimates of this paper. We begin by introducing some notation. For each integer k, define the efficient differencing operator $\Delta^*_{1,k}$ by

$$\Delta^*_{1,k}(f(x);h;m) = f(x+hm^k) - f(x),$$

and define $\Delta^*_{j,\mathbf{K}}$ successively by

$$\begin{aligned}\Delta^*_{j+1,K_1,\dots,K_{j+1}}&(f(x);h_1,\dots,h_{j+1};m_1,\dots,m_{j+1}) \\ &= \Delta^*_{1,K_{j+1}}\left(\Delta^*_{j,K_1,\dots,K_j}(f(x);h_1,\dots,h_j;m_1,\dots,m_j);h_{j+1};m_{j+1}\right).\end{aligned}$$

Also, adopt the convention that

$$\Delta^*_{0,k}(f(x);h;m) = f(x).$$

When $0 \le j \le k_1$, let $K_1,\dots,K_j$ be fixed exponents associated with the efficient differencing process, and let

$$\Psi_{i,j} = \Psi^{(\mathbf{K})}_{i,j}(z;h_1,\dots,h_j;m_1,\dots,m_j)$$

be defined by

$$\Psi_{i,j} = \Delta^*_{j,\mathbf{K}}(z^{k_i};h_1,\dots,h_j;m_1,\dots,m_j).$$

We consider the effect of substituting $\Psi_{i,j}(z;\mathbf{h};\mathbf{m})$ for $\Psi_i(z,\mathbf{c})$ in Lemmata 3.1 and 3.2. We require some further notation in order to discuss the consequences of this substitution. When $k_1,\dots,k_t$ satisfy (1.2), define $\widetilde{r}_J$ by

$$\widetilde{r}_J = \operatorname{card}\{i \in [1,t] : k_i > J\}.$$

We note that $\Psi_{i,j}$ has positive degree whenever $1 \le i \le \widetilde{r}_j$, and moreover from (1.2) it follows easily that for each J one has $\widetilde{r}_J \ge t - J$. Suppose that K_j $(1 \le j \le k_1)$ are positive integers with

$$\sum_{j=1}^{k_1} K_j^{-1} < 1,$$

and that r_j $(0 \le j < k_1)$ are positive integers with $r_j \le \widetilde{r}_j$. Take $\phi_j = \phi_j(\mathbf{r},s,\mathbf{k},\mathbf{K})$ $(1 \le j \le k_1)$ to be real numbers with $0 \le \phi_j \le 1/K_j$ to be determined later. Write

$$M_j = P^{\phi_j}, \quad H_j = PM_j^{-K_j} \quad \text{and} \quad Q_j = P(M_1\dots M_j)^{-1} \quad (1 \le j \le k_1), \tag{4.1}$$

and further, for the sake of convenience, write

$$\widetilde{M}_j = \prod_{i=1}^{j} M_i \quad \text{and} \quad \widetilde{H}_j = \prod_{i=1}^{j} H_i. \tag{4.2}$$

Substitute the conditions

$$M_i < m_i \le M_i R \quad \text{and} \quad 1 \le h_i \le H_i \quad (1 \le i \le j), \tag{4.3}$$

for the conditions $C'_j < c_j \le C_j$ $(1 \le j \le u)$ in (3.3), and put

$$D_i(\mathbf{m}) = \prod_{l=1}^{j} m_l^{k_i} \quad (1 \le i \le t). \tag{4.4}$$

Write $S_{s,r}(P,Q,R;\boldsymbol{\Psi}_j)$ for $S^{(\mathbf{k})}_{s,r}(P,Q,R;\boldsymbol{\Psi};\mathbf{C},\mathbf{C}';\mathbf{D})$, and do likewise for $\widetilde{S}_{s,r}$. Also, put $\theta = \phi_{j+1}$, and adopt similar conventions to those above for $T_{s,r}$ and $\widetilde{T}_{s,r,k}$. Finally, write $J_r(\mathbf{z};\mathbf{h};\mathbf{m})$ for $J_r(\mathbf{z};\mathbf{c})$.

Lemma 4.1. *There exists a positive number* $\eta = \eta(\varepsilon, s, \mathbf{k})$ *such that whenever* $\exp\left((\log\log P)^2\right) < R \le P^\eta$*, and* $0 \le j < k_1$*, one has*

$$S_{s,r_j}(P, Q_j, R; \mathbf{\Psi}_j) \ll P^\varepsilon M_{j+1}^{2s-1+\omega(\mathbf{k}; K_{j+1}, r_j, r_j)} \widetilde{T}_{s,r_j,K_{j+1}}(P, Q_j, R; \phi_{j+1}; \mathbf{\Psi}_j).$$

Proof. Write, for the sake of convenience,

$$\phi = \phi_{j+1}, \quad k = K_{j+1}, \quad r = r_j \quad \text{and} \quad \omega = \omega(\mathbf{k}; K_{j+1}, r_j, r_j).$$

Then the condition $r \le \widetilde{r}_j$ enables us to apply Lemmata 3.1 and 3.2 to deduce that

$$\begin{aligned} S_{s,r}(P, Q_j, R; \mathbf{\Psi}_j) &\ll P^\varepsilon Q_j M_{j+1} \widetilde{S}_{s-1,r}(P, Q_j, R; \mathbf{\Psi}_j) \\ &\quad + P^\varepsilon M_{j+1}^{2s-1+\omega} \widetilde{T}_{s,r,k}(P, Q_j, R; \phi; \mathbf{\Psi}_j). \end{aligned} \tag{4.5}$$

We show inductively that for $s = 1, 2, \ldots$, one has

$$Q_j M_{j+1} \widetilde{S}_{s-1,r}(P, Q_j, R; \mathbf{\Psi}_j) \ll P^\varepsilon M_{j+1}^{2s-1+\omega} \widetilde{T}_{s,r,k}(P, Q_j, R; \phi; \mathbf{\Psi}_j). \tag{4.6}$$

This will complete the proof of the lemma.

We start by observing that for $s = 1$, the quantity $\widetilde{S}_{s-1,r}(P, Q_j, R; \mathbf{\Psi}_j)$ is bounded above by the number of solutions of the system

$$\sum_{n=1}^{r} \eta_n \left(\Psi_{i,j}(z_n; \mathbf{h}; \mathbf{m}) - \Psi_{i,j}(w_n; \mathbf{h}; \mathbf{m})\right) = 0 \quad (1 \le i \le r), \tag{4.7}$$

with $J_r(\mathbf{z}; \mathbf{h}; \mathbf{m}) \ne 0$, $J_r(\mathbf{w}; \mathbf{h}; \mathbf{m}) \ne 0$, and

$$1 \le z_n, w_n \le P, \quad \eta_n \in \{+1, -1\} \quad (1 \le n \le r),$$
$$M_i < m_i \le M_i R, \quad 1 \le h_i \le H_i \quad (1 \le i \le j).$$

There are $O(P^{r+\varepsilon} \widetilde{H}_j \widetilde{M}_j)$ permissible choices for $\mathbf{w}$, $\mathbf{h}$ and $\mathbf{m}$. Fix any one such choice, and consider the system of equations (4.7) in $\mathbf{z}$. Since $J_r(\mathbf{z}; \mathbf{h}; \mathbf{m}) \ne 0$, it follows from the Inverse Function Theorem that there are $O_{\mathbf{k}}(1)$ solutions $\mathbf{z}$ of (4.7). Consequently, when $s = 1$ one has

$$\widetilde{S}_{s-1,r}(P, Q_j, R; \mathbf{\Psi}_j) \ll P^{r+\varepsilon} \widetilde{H}_j \widetilde{M}_j.$$

But when $R > \exp\left((\log\log Q)^2\right)$ a standard estimate yields

$$\operatorname{card}(\mathcal{A}(Q, R)) \gg Q^{1-\varepsilon}. \tag{4.8}$$

Thus, on considering the diagonal solutions of the system (3.6) with $s = 1$, one finds that

$$\widetilde{T}_{s,r,k}(P, Q_j, R; \phi; \mathbf{\Psi}_j) \gg P^r \widetilde{H}_j \widetilde{M}_j M_{j+1} Q_{j+1}^{1-\varepsilon},$$

whence (4.6) follows in the case $s = 1$.

Now suppose that $s > 1$, and that the inductive hypothesis holds for $S < s$. On considering the underlying diophantine equations, for each s we obtain

$$\widetilde{S}_{s,r}(P, Q_j, R; \mathbf{\Psi}_j) \le S_{s,r}(P, Q_j, R; \mathbf{\Psi}_j).$$

Then by combining the inductive hypothesis with (4.5) one deduces that

$$\widetilde{S}_{s-1,r}(P, Q_j, R; \mathbf{\Psi}_j) \ll P^\varepsilon M_{j+1}^{2s-3+\omega} \widetilde{T}_{s-1,r,k}(P, Q_j, R; \phi; \mathbf{\Psi}_j). \tag{4.9}$$

But by considering solutions of (3.6) in which $u_s = v_s$, it follows from (4.8) that

$$\widetilde{T}_{s,r,k}(P, Q_j, R; \phi; \mathbf{\Psi}_j) \gg (Q_j/M_{j+1})^{1-\varepsilon} \widetilde{T}_{s-1,r,k}(P, Q_j, R; \phi; \mathbf{\Psi}_j), \tag{4.10}$$

and so (4.6) follows from (4.9) and (4.10). Thus (4.6) holds for each positive integer s, and the proof of the lemma is complete.

Lemma 4.1 permits us to relate $S_{s,r}(P, Q_j, R; \mathbf{\Psi}_j)$ to $\widetilde{T}_{s,r,k}(P, Q_j, R; \phi; \mathbf{\Psi}_j)$. We now consider the simplest method of relating $\widetilde{T}_{s,r,k}(P, Q_j, R; \phi; \mathbf{\Psi}_j)$ to expressions of the form $S_{s,w}(P, Q_{j+1}, R; \mathbf{\Psi}_{j+1})$. We thereby obtain estimates for $S_{s,r}(P, P, R; \mathbf{\Psi}_0)$, and hence for $S_{s+r}(P, R)$.

Lemma 4.2. *Let j be an integer with $0 \le j < k_1$. Write*

$$r = r_j, \quad w = r_{j+1}, \quad k = K_{j+1},$$

and

$$\widetilde{T}_{s,r,k} = \widetilde{T}_{s,r_j,K_{j+1}}(P, Q_j, R; \phi_{j+1}; \mathbf{\Psi}_j).$$

Suppose that r and w satisfy $1 \le r \le 2w$. Then

$$\begin{aligned}\widetilde{T}_{s,r,k} \ll & P^{r+\varepsilon}\widetilde{H}_j\widetilde{M}_{j+1}S_s(Q_{j+1}, R)\\ & + P^{\varepsilon}H_{j+1}^{r-1}\left(\widetilde{H}_{j+1}\widetilde{M}_{j+1}S_s(Q_{j+1}, R)\right)^{1-\frac{r}{2w}}\left(S_{s,w}(P, Q_{j+1}, R; \mathbf{\Psi}_{j+1})\right)^{\frac{r}{2w}}.\end{aligned}$$

Proof. We consider the system (3.6) with $\theta = \phi_{j+1}$, and recall the definition of $\widetilde{T}_{s,r,k}$. Write

$$\mathcal{L}_{a,d}(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m}) = \sum_{\substack{1\le z\le P\\ z\equiv a \pmod d}} e\left(\alpha_1\Psi_{1,j}(z; \mathbf{h}; \mathbf{m}) + \cdots + \alpha_t\Psi_{t,j}(z; \mathbf{h}; \mathbf{m})\right),$$

and

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

Write also

$$g_q(\boldsymbol{\alpha}; \mathbf{m}) = \sum_{x\in\mathcal{A}(Q_{j+1},R)} e\left(\sum_{i=1}^{t}\alpha_i D_i(\mathbf{m})(qx)^{k_i}\right).$$

Then on considering the underlying diophantine equations, we find that

$$\widetilde{T}_{s,r,k} \ll \sum_{\mathbf{h},\mathbf{m}}\ \sum_{M_{j+1}<q\le M_{j+1}R} \int_{\mathbb{T}^t} \mathcal{K}_{q^k}(\boldsymbol{\alpha}; \mathbf{h}, \mathbf{m})^r \left|g_q(\boldsymbol{\alpha}; \mathbf{m})\right|^{2s} d\boldsymbol{\alpha} \ll \widetilde{T}_{s,r,k}, \tag{4.11}$$

where the summation is over $\mathbf{h}$ and $\mathbf{m}$ satisfying (4.3).

Let U_0 denote the number of solutions of the system (3.6) counted by $\widetilde{T}_{s,r,k}$ in which $z_n = w_n$ for at least one n with $1 \le n \le r$, and let U_1 denote the corresponding number of solutions with $z_n \ne w_n$ $(1 \le n \le r)$. Then

$$\widetilde{T}_{s,r,k} = U_0 + U_1.$$

We divide into cases.

(i) Suppose that $U_0 \ge U_1$, so that $\widetilde{T}_{s,r,k} \ll U_0$. Then by considering the underlying diophantine equations,

$$U_0 \ll P \sum_{\mathbf{h},\mathbf{m}}\ \sum_{M_{j+1}<q\le M_{j+1}R} \int_{\mathbb{T}^t} \mathcal{K}_{q^k}(\boldsymbol{\alpha}; \mathbf{h}, \mathbf{m})^{r-1} \left|g_q(\boldsymbol{\alpha}; \mathbf{m})\right|^{2s} d\boldsymbol{\alpha}.$$

An application of Hölder's inequality gives

$$\begin{aligned}\widetilde{T}_{s,r,k} \ll P & \left(\sum_{\mathbf{h},\mathbf{m}}\ \sum_{M_{j+1}<q\le M_{j+1}R} \int_{\mathbb{T}^t} \mathcal{K}_{q^k}(\boldsymbol{\alpha}; \mathbf{h}, \mathbf{m})^r \left|g_q(\boldsymbol{\alpha}; \mathbf{m})\right|^{2s} d\boldsymbol{\alpha}\right)^{1-1/r}\\ & \times \left(\sum_{\mathbf{h},\mathbf{m}}\ \sum_{M_{j+1}<q\le M_{j+1}R} \int_{\mathbb{T}^t} \left|g_q(\boldsymbol{\alpha}; \mathbf{m})\right|^{2s} d\boldsymbol{\alpha}\right)^{1/r},\end{aligned}$$

whence by (4.11),

$$\widetilde{T}_{s,r,k} \ll P^{1+\varepsilon}\left(\widetilde{T}_{s,r,k}\right)^{1-1/r}\left(\widetilde{H}_j\widetilde{M}_{j+1}S_s(Q_{j+1}, R)\right)^{1/r}.$$

Then
$$\widetilde{T}_{s,r,k} \ll P^{r+\varepsilon}\widetilde{H}_j\widetilde{M}_{j+1}S_s(Q_{j+1},R),$$
which completes the proof of the lemma in the first case.

(ii) Suppose that $U_1 > U_0$, so that $\widetilde{T}_{s,r,k} \ll U_1$. We start by noting that for each solution of (3.6) counted by U_1 we have
$$z_n \equiv w_n \pmod{q^k} \quad \text{and} \quad z_n \neq w_n \quad (1 \le n \le r).$$
Then for each n with $1 \le n \le r$, there exists an integer g_n with $1 \le |g_n| \le H_{j+1}$ such that
$$w_n = z_n + g_n q^k. \tag{4.12}$$
On substituting (4.12) into (3.6), we deduce that
$$U_1 \le \sum_{\boldsymbol{\eta}\in\{+1,-1\}^r} U_2(\boldsymbol{\eta}),$$
where $U_2(\boldsymbol{\eta})$ denotes the number of solutions of the system
$$\sum_{l=1}^r \eta_l \Psi_{i,j+1}(z_l;\mathbf{h},g_l;\mathbf{m},q) + D_i(\mathbf{m})q^{k_i}\sum_{m=1}^s (u_m^{k_i} - v_m^{k_i}) = 0 \quad (1 \le i \le t),$$
with $\mathbf{z},\mathbf{u},\mathbf{v},\mathbf{h},\mathbf{m}$ satisfying (3.4), (3.7), (4.3),
$$1 \le g_l \le H_{j+1} \quad (1 \le l \le r) \quad \text{and} \quad M_{j+1} < q \le M_{j+1}R.$$

Write
$$G(\boldsymbol{\alpha};g,q) = \sum_{1\le z\le P} e(\alpha_1\Psi_{1,j+1}(z;\mathbf{h},g;\mathbf{m},q) + \dots + \alpha_t\Psi_{t,j+1}(z;\mathbf{h},g;\mathbf{m},q)).$$
Then on considering the underlying diophantine equations, we deduce that
$$U_2(\boldsymbol{\eta}) \ll \sum_{\mathbf{h},\mathbf{m}}\ \sum_{M_{j+1}<q\le M_{j+1}R} \int_{\mathbb{T}^t} \Bigl| \sum_{1\le g\le H_{j+1}} G(\boldsymbol{\alpha};g,q)\Bigr|^r |g_q(\boldsymbol{\alpha};\mathbf{m})|^{2s}\, d\boldsymbol{\alpha}. \tag{4.13}$$
Repeated application of Hölder's inequality to (4.13) reveals that
$$U_2(\boldsymbol{\eta}) \ll H_{j+1}^{r-1} V_1^{r/2w} V_2^{1-r/2w}, \tag{4.14}$$
where
$$V_1 = \sum_{\mathbf{h},\mathbf{m}}\ \sum_{1\le g\le H_{j+1}}\ \sum_{M_{j+1}<q\le M_{j+1}R} \int_{\mathbb{T}^t} \bigl|G(\boldsymbol{\alpha};g,q)^{2w} g_q(\boldsymbol{\alpha};\mathbf{m})^{2s}\bigr|\, d\boldsymbol{\alpha},$$
and
$$V_2 = \sum_{\mathbf{h},\mathbf{m}}\ \sum_{1\le g\le H_{j+1}}\ \sum_{M_{j+1}<q\le M_{j+1}R} \int_{\mathbb{T}^t} |g_q(\boldsymbol{\alpha};\mathbf{m})|^{2s}\, d\boldsymbol{\alpha}.$$
But on considering the underlying diophantine equations, one has
$$V_1 \ll S_{s,w}(P,Q_{j+1},R;\boldsymbol{\Psi}_{j+1}),$$
and
$$V_2 \ll P^{\varepsilon}\widetilde{H}_{j+1}\widetilde{M}_{j+1}S_s(Q_{j+1},R),$$
and thus the proof of the lemma is completed on recalling (4.14).

To conclude this section, we note that it is plainly possible to apply Hölder's inequality a little differently in the final stages of the proof of Lemma 4.2, so that $U_2(\boldsymbol{\eta})$ is bounded in terms of $S_{s_1,w}(P,Q_{j+1},R;\boldsymbol{\Psi}_{j+1})$ and $S_{s_2}(Q_{j+1},R)$ for suitable $s_1 \neq s_2$. Thereby one can employ iterative schemes analogous to those used in [20] in order to improve on the bounds we ultimately obtain. Since, however, these improvements are likely to be significant only for smaller $\mathbf{k}$, we will not pursue this idea further.

5. Mean value estimates based on single differencing

We are now in a position to exploit the machinery developed in the previous two sections. This section will be devoted to proving Theorem 1. We first record some notation. Let k be an integer with $1 \le k \le k_1$, and let θ be a parameter with $0 < \theta \le 1/k$ to be chosen later. Write

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

Recall (1.3), and write $f(\boldsymbol{\alpha}; P, R)$ for $f_{\mathbf{k}}(\boldsymbol{\alpha}; P, R)$. Define also

$$F(\boldsymbol{\alpha}) = \sum_{1 \le z \le P} e(\alpha_1 z^{k_1} + \dots + \alpha_t z^{k_t}), \tag{5.1}$$

$$G(\boldsymbol{\alpha}; q) = \sum_{1 \le h \le H} \sum_{1 \le z \le P} e\left(\sum_{i=1}^{t} \alpha_i \Psi_{i,1}(z; h; q) \right),$$

and

$$g_q(\boldsymbol{\alpha}; Q, R) = \sum_{x \in \mathcal{A}(Q,R)} e\left(\sum_{i=1}^{t} \alpha_i (xq)^{k_i} \right).$$

Finally, define the mean values

$$\mathcal{M}_{s,r}(P, R) = \sum_{M < q \le MR} \int_{\mathbb{T}^t} \left| G(\boldsymbol{\alpha}; q)^r g_q(\boldsymbol{\alpha}; Q, R)^{2s} \right| d\boldsymbol{\alpha},$$

and

$$\mathcal{S}_{s,r}(P, R) = \int_{\mathbb{T}^t} \left| F(\boldsymbol{\alpha})^{2r} f(\boldsymbol{\alpha}; P, R)^{2s} \right| d\boldsymbol{\alpha}.$$

It is convenient to incorporate in the following lemma the key features of the general efficient differencing process described in §§3 and 4.

Lemma 5.1. *Suppose that $1 \le r \le t$, $1 \le k \le k_1$ and $0 < \theta \le 1/k_1$. Then*

$$\mathcal{S}_{s,r}(P, R) \ll P^{\varepsilon} M^{2s-1+\omega(\mathbf{k}; k, r, r)} \left(P^r M S_s(Q, R) + \mathcal{M}_{s,r}(P, R) \right).$$

Proof. On taking $\Psi_i(z; \mathbf{c}) = z^{k_i}$ $(1 \le i \le t)$, and considering the underlying diophantine equations, it follows from the definition of $S_{s,r}^{(\mathbf{k})}$ that $\mathcal{S}_{s,r}(P, R) = S_{s,r}^{(\mathbf{k})}(P, P, R; \boldsymbol{\Psi}_0)$. Then by Lemma 4.1 with $j = 0$,

$$\mathcal{S}_{s,r}(P, R) \ll P^{\varepsilon} M^{2s-1+\omega(\mathbf{k}; k, r, r)} \widetilde{T}_{s,r,k}(P, P, R; \theta; \boldsymbol{\Psi}_0). \tag{5.2}$$

Moreover, recalling our change in notation in this section, it follows from the argument of the proof of Lemma 4.2 (see equation (4.11) together with part (i) and the argument of part (ii) leading to (4.13)) that

$$\widetilde{T}_{s,r,k}(P, P, R; \theta; \boldsymbol{\Psi}_0) \ll P^{r+\varepsilon} M S_s(Q, R) + \mathcal{M}_{s,r}(P, R). \tag{5.3}$$

The lemma follows on combining (5.2) and (5.3).

The next lemma shows how a permissible exponent $\lambda_{s+t,\mathbf{k}}$ can be obtained from a permissible exponent $\lambda_{s,\mathbf{k}}$. Before stating this lemma, we equip ourselves with a further definition. We shall say that the exponent $\Delta_{s,\mathbf{k}}$ is *admissible* whenever it has the property that

$$S_s(P, R) \ll P^{\lambda_{s,\mathbf{k}}+\varepsilon},$$

with $\lambda_{s,\mathbf{k}} = 2s - \sum_{i=1}^{t} k_i + \Delta_{s,\mathbf{k}}$. Thus $\Delta_{s,\mathbf{k}}$ is admissible whenever $\lambda_{s,\mathbf{k}}$ is permissible.

Theorem 5.2. *Suppose that $\Delta_{s,\mathbf{k}}$ is an admissible exponent. Then the exponent $\Delta_{s+t,\mathbf{k}}$ is admissible, where $\Delta_{s+t,\mathbf{k}} = \Delta_{s,\mathbf{k}}(1-1/k_1)$.*

Proof. We apply Lemma 5.1 with $k = k_1$, $\theta = 1/k_1$ and $r = t$, noting that $S_{s,t}(P,P,R;\boldsymbol{\Psi}_0) = \mathcal{S}_{s,t}(P,R)$. Since with the given parameters we have $H = 1$, we obtain by a trivial estimate

$$\mathcal{M}_{s,t}(P,R) \ll P^t \sum_{M<q\le MR} \int_{\mathbb{T}^t} |g_q(\boldsymbol{\alpha};Q,R)|^{2s}\, d\boldsymbol{\alpha} \ll P^{t+\varepsilon} M S_s(Q,R).$$

We therefore deduce from Lemma 5.1 that

$$S_{s,t}(P,P,R;\boldsymbol{\Psi}_0) \ll P^{t+\varepsilon} M^{2s+\omega(\mathbf{k};k_1,t,t)} S_s(Q,R). \tag{5.4}$$

We suppose that Δ_s is admissible, and write $\lambda_s = 2s - \sum_{i=1}^t k_i + \Delta_s$. Then on recalling (2.2) and considering the underlying diophantine equations,

$$S_{s+t}(P,R) \ll S_{s,t}(P,P,R;\boldsymbol{\Psi}_0) \ll P^{\lambda_{s+t}+\varepsilon},$$

where

$$\begin{aligned}\lambda_{s+t} &= t + \left(2s + tk_1 - \sum_{i=1}^t k_i\right)\theta + \lambda_s(1-\theta)\\ &= 2s + 2t - \sum_{i=1}^t k_i + \Delta_s(1-1/k_1).\end{aligned}$$

It follows that the exponent Δ_{s+t} is admissible, where $\Delta_{s+t} = \Delta_s(1-1/k_1)$, and this completes the proof of the lemma.

The proof of Theorem 1. By considering the underlying diophantine equations, it follows from Theorem 1 of [25] that $S_{t+1}(P,R) \ll P^{t+1+\varepsilon}$. Then for each r with $1 \le r \le t+1$, the exponent $\Delta_r = \sum_{i=1}^t k_i - r$ is admissible. Consequently an inductive argument reveals that the exponent

$$\Delta_{lt+r} = \left(\sum_{i=1}^t k_i - r\right)(1-1/k_1)^l$$

is admissible for each $l \in \mathbb{N}$. Theorem 1 follows immediately.

We remark that the methods of this section suffice to establish quasi-diagonal behaviour in the mean value $S_s^{(\mathbf{k})}(P,R)$ (see §8 of [27] for a sketch of the necessary argument).

6. Estimates arising from repeated efficient differencing

Mean value estimates based on repeated efficient differencing are usually sharper, though more complicated to state, than those arising from a single efficient differencing process. In Theorem 6.1 we describe the permissible exponents which can be obtained by applying Lemmata 4.1 and 4.2 successively in a relatively simple manner. Our strategy is to use Lemma 4.1 to bound $S_{s,r_j}(P,Q_j,R;\boldsymbol{\Psi}_j)$ in terms of $\widetilde{T}_{s,r_j,K_{j+1}}(P,Q_j,R;\phi_{j+1};\boldsymbol{\Psi}_j)$, and then to use Lemma 4.2 to bound the latter in terms of $S_s(Q_{j+1},R)$ and $S_{s,r_{j+1}}(P,Q_{j+1},R;\boldsymbol{\Psi}_{j+1})$. We can repeat this process non-trivially as many as k_1 times, optimising the parameters ϕ_j, r_j and K_j when we terminate the process.

Theorem 6.1. *Let $k_1,\dots,k_t$ be integers satisfying* (1.2). *Suppose that u is a positive integer, and that $\lambda_{u,\mathbf{k}}$ is a permissible exponent satisfying*

$$2u - \sum_{i=1}^t k_i < \lambda_{u,\mathbf{k}} \le 2u. \tag{6.1}$$

Write

$$\Delta_u = \lambda_{u,\mathbf{k}} - 2u + \sum_{i=1}^t k_i. \tag{6.2}$$

For $s=u+lt$ $(l\in\mathbb{N})$, define the real numbers Δ_s, θ_s and $\phi(j,s,J)$ recursively as follows. For $l\ge 1$ and $j=1,\dots,k_1$, put $\phi(j,s,j)=1/k_1$, and evaluate $\phi(j,s,J-1)$ successively for $J=j,\dots,2$ by

$$\phi^*(j,s,J-1)=\frac{1}{2k_1}+\left(\frac{1}{2}+\frac{\Omega_{J-1}-\Delta_{s-t}}{2k_1\widetilde{r}_{J-1}}\right)\phi(j,s,J), \tag{6.3}$$

where (recalling the definition of $\widetilde{r}_J$ from §4)

$$\Omega_J=\sum_{l>\widetilde{r}_J}k_l, \tag{6.4}$$

and

$$\phi(j,s,J-1)=\min\{1/k_1,\phi^*(j,s,J-1)\}. \tag{6.5}$$

Finally, set

$$\theta_s=\min_{1\le j\le k_1}\phi(j,s,1), \tag{6.6}$$

$$\Delta_s=\Delta_{s-t}(1-\theta_s)+t(k_1\theta_s-1), \tag{6.7}$$

and

$$\lambda_s=2s-\sum_{i=1}^{t}k_i+\Delta_s. \tag{6.8}$$

Then λ_s is a permissible exponent for $s=u+lt$ $(l\in\mathbb{N})$.

The conclusion of Theorem 6.1 shows that for the exponents λ_s defined by (6.8), given $\varepsilon>0$, there is a positive number $\eta=\eta(\varepsilon,s,\mathbf{k})$ such that whenever $R\le P^\eta$, one has

$$\int_{\mathbb{T}^t}|f_{\mathbf{k}}(\boldsymbol{\alpha};P,R)|^{2s}\,d\boldsymbol{\alpha}\ll P^{\lambda_s+\varepsilon}.$$

We note that the argument which leads to Theorem 6.1 in fact yields a stronger conclusion. Under the same hypotheses, when $s>t$ one has

$$\int_{\mathbb{T}^t}\left|F(\boldsymbol{\alpha})^{2t}f_{\mathbf{k}}(\boldsymbol{\alpha};P,R)^{2s-2t}\right|d\boldsymbol{\alpha}\ll P^{\lambda_s+\varepsilon},$$

where $F(\boldsymbol{\alpha})$ is defined by (5.1).

Proof of Theorem 6.1. Before starting the proof of the theorem, we make some comments concerning the variables in its statement. Notice first that for each j, s and J, we have $\phi(j,s,J)\le 1/k_1$. Therefore $\theta_s\le 1/k_1$, and hence a trivial induction shows that

$$\Delta_s\le\max\{0,\Delta_{s-t}\}\le\sum_{i=1}^{t}k_i\le\widetilde{r}_Jk_1+\sum_{l>\widetilde{r}_J}k_l\quad(1\le J\le k_1).$$

Then (6.3) yields positive values for the ϕ^*, and hence also for the ϕ, and thus $\theta_s>0$.

We prove the theorem by induction, starting with the assumption that the exponents λ_{u+mt} defined by (6.8) are permissible when $0\le m\le l$. We write $s=u+lt$, and aim to prove that the exponent λ_{s+t} defined by (6.8) is permissible. Let j be the least integer with $1\le j\le k_1$ such that $\theta_{s+t}=\phi(j,s+t,1)$. For $J=1,\dots,j$ define $\phi_J=\phi(j,s+t,J)$ as in the statement of the theorem. Then if $\phi_J=1/k_1$ for some $J<j$, we have $\phi(j,s+t,J)=\phi(J,s+t,J)$, and one finds successively that $\phi(j,s+t,r)=\phi(J,s+t,r)$ for $r=J,J-1,\dots,1$, contradicting the minimality of j. Thus $\phi_J<1/k_1$ for $J<j$. We adopt the notation of writing

$$M_i=P^{\phi_i},\quad H_i=PM_i^{-k_1},\quad Q_i=P(M_1\dots M_i)^{-1}\quad(1\le i\le j),$$

and adopt the convention that $Q_0=P$.

We first prove inductively that for $J=j-1,j-2,\dots,0$, one has

$$\widetilde{T}_{s,\widetilde{r}_J,k_1}(P,Q_J,R;\phi_{J+1};\mathbf{\Psi}_J)\ll P^{\widetilde{r}_J+\varepsilon}\widetilde{H}_J\widetilde{M}_{J+1}Q_{J+1}^{\lambda},\tag{6.9}$$

where for convenience we write $\lambda=2s-\sum_{i=1}^t k_i+\Delta_s$.

On applying Lemma 4.2 with j replaced by $j-1$, and $r_J=\widetilde{r}_J$ $(J=j-1,j)$, we obtain

$$\widetilde{T}_{s,\widetilde{r}_{j-1},k_1}(P,Q_{j-1},R;\phi_j;\mathbf{\Psi}_{j-1})\ll P^{\varepsilon}(U_1+U_2),$$

where

$$U_1=P^{\widetilde{r}_{j-1}}\widetilde{H}_{j-1}\widetilde{M}_jS_s(Q_j,R),$$

$$U_2=H_j^{\widetilde{r}_{j-1}-1}\left(\widetilde{H}_j\widetilde{M}_jS_s(Q_j,R)\right)^{1-\gamma}\left(S_{s,\widetilde{r}_j}(P,Q_j,R;\mathbf{\Psi}_j)\right)^{\gamma},$$

and we have written $\gamma=\widetilde{r}_{j-1}/(2\widetilde{r}_j)$. Then on noting the trivial estimate

$$S_{s,\widetilde{r}_j}(P,Q_j,R;\mathbf{\Psi}_j)\ll P^{2\widetilde{r}_j+\varepsilon}\widetilde{H}_j\widetilde{M}_jS_s(Q_j,R),$$

and observing that since $\phi(j,s+t,j)=1/k_1$ one has $H_j=1$, we deduce that (6.9) holds in the case $J=j-1$.

We now suppose that (6.9) holds for J and deduce the corresponding inequality for $J-1$, noting that since we have just established (6.9) when $J=j-1$, we may suppose that $1\le J\le j-1$. It follows from Lemma 4.1 together with the hypothesis that (6.9) holds for J that

$$S_{s,\widetilde{r}_J}(P,Q_J,R;\mathbf{\Psi}_J)\ll P^{\varepsilon}M_{J+1}^{2s+\omega(\mathbf{k};k_1,\widetilde{r}_J,\widetilde{r}_J)}P^{\widetilde{r}_J}\widetilde{H}_J\widetilde{M}_JQ_{J+1}^{\lambda}.\tag{6.10}$$

Also by Lemma 4.2 with j replaced by $J-1$, and $r_l=\widetilde{r}_l$ $(l=J-1,J)$, we obtain

$$\widetilde{T}_{s,\widetilde{r}_{J-1},k_1}(P,Q_{J-1},R;\phi_J;\mathbf{\Psi}_{J-1})\ll P^{\varepsilon}(V_1+V_2),\tag{6.11}$$

where

$$V_1=P^{\widetilde{r}_{J-1}}\widetilde{H}_{J-1}\widetilde{M}_JS_s(Q_J,R),\tag{6.12}$$

$$V_2=H_J^{\widetilde{r}_{J-1}-1}\left(\widetilde{H}_J\widetilde{M}_JS_s(Q_J,R)\right)^{1-\gamma'}\left(S_{s,\widetilde{r}_J}(P,Q_J,R;\mathbf{\Psi}_J)\right)^{\gamma'},\tag{6.13}$$

and we have written $\gamma'=\widetilde{r}_{J-1}/(2\widetilde{r}_J)$. We have assumed that $\phi_J<1/k_1$ for $J<j$, and hence that $\phi_J=\phi^*(j,s+t,J)$. Then since $\lambda=2s-\sum_{i=1}^t k_i+\Delta_s$, we deduce from (6.3) that

$$\begin{aligned}(2s+\omega(\mathbf{k};k_1,\widetilde{r}_J,\widetilde{r}_J))\phi_{J+1}-\lambda\phi_{J+1}&=(\widetilde{r}_Jk_1+\Omega_J-\Delta_s)\phi(j,s+t,J+1)\\&=2\widetilde{r}_Jk_1\phi_J-\widetilde{r}_J.\end{aligned}\tag{6.14}$$

Then it follows from (6.10), (6.13) and (6.14) that

$$V_2\ll P^{\widetilde{r}_{J-1}+\varepsilon}\widetilde{H}_{J-1}\widetilde{M}_JQ_J^{\lambda}.$$

Consequently, by (6.11) and (6.12),

$$\widetilde{T}_{s,\widetilde{r}_{J-1},k_1}(P,Q_{J-1},R;\phi_J;\mathbf{\Psi}_{J-1})\ll P^{\widetilde{r}_{J-1}+\varepsilon}\widetilde{H}_{J-1}\widetilde{M}_JQ_J^{\lambda},$$

whence (6.9) follows with $J-1$ replacing J, and our second assertion holds for $J=0,\dots,j-1$.

On noting that $\widetilde{r}_0=t$, we next observe that

$$\widetilde{T}_{s,t,k_1}(P,Q_0,R;\phi_1;\mathbf{\Psi}_0)\ll P^{t+\varepsilon}M_1Q_1^{\lambda},$$

so that by Lemma 4.1,

$$S_{s,t}(P,Q_0,R;\mathbf{\Psi}_0)\ll P^{t+\varepsilon}M_1^{2s+\omega(\mathbf{k};k_1,t,t)}Q_1^{\lambda}.$$

Then

$$S_{s+t}(P,R)\ll S_{s,t}(P,P,R;\mathbf{\Psi}_0)\ll P^{\lambda'+\varepsilon},$$

where

$$\begin{aligned}\lambda'&=\lambda(1-\theta_{s+t})+t+(2s+\omega(\mathbf{k};k_1,t,t))\theta_{s+t}\\&=2(s+t)-\sum_{i=1}^t k_i+\Delta_{s+t},\end{aligned}$$

and

$$\Delta_{s+t}=\Delta_s(1-\theta_{s+t})+t(k_1\theta_{s+t}-1).$$

Thus our inductive hypothesis follows with $s+t$ in place of s, and this completes the proof of the theorem.

We exploit Theorem 6.1 to obtain Theorems 2 and 3 through the use of the following lemma.

Lemma 6.2. *Let $k_1,\dots,k_t$ be integers satisfying* (1.2) *with k_1 sufficiently large. Suppose that $s>t$, and that $\Delta_{s-t,\mathbf{k}}$ is an admissible exponent satisfying $\Delta_{s-t,\mathbf{k}}>(\log k_1)^2$. Write $\delta_{s-t}=\Delta_{s-t,\mathbf{k}}/(tk_1)$, and define δ_s to be the unique positive solution of the equation*

$$\delta_s+\log\delta_s=\delta_{s-t}+\log\delta_{s-t}-\frac{2}{k_1}+\frac{2}{k_1(\log k_1)^{3/2}}.$$

Then the exponent $\Delta_{s,\mathbf{k}}=tk_1\delta_s$ is admissible.

Proof. Our argument is modelled on the proof of Theorem 2.1 of [26], as modified in [3]. We adopt the same notation as in the statement of the proof of Theorem 6.1, and write $k=k_1$. We start with some simplifying comments. First note that since $0\le\theta_s\le 1/k$ for each s, then whenever $0<\Delta_{s-t}\le tk$ the equation (6.7) implies that

$$\Delta_s=\Delta_{s-t}(1-1/k)+(tk-\Delta_{s-t})(\theta_s-1/k)\le\Delta_{s-t}(1-1/k).$$

Then for each fixed r it follows that Δ_{lt+r} is a decreasing function of l for $l\ge1$. We may consequently suppose that $0\le\Delta_s\le tk$ for every admissible exponent Δ_s, whence $0\le\delta_s\le1$. Moreover, on noting that $\delta+\log\delta$ is an increasing function of δ when $\delta>0$, in order to establish the lemma it suffices to prove that the exponent $\Delta_s=kt\delta_s$ is admissible, where δ_s is any positive number satisfying

$$\delta_s+\log\delta_s\le\delta_{s-t}+\log\delta_{s-t}-\frac{2}{k}+\frac{2}{k(\log k)^{3/2}}. \tag{6.15}$$

Suppose that s is an integer with $s>t$, and that Δ_{s-t} is an admissible exponent satisfying $\Delta_{s-t}>(\log k)^2$. We apply Theorem 6.1 with $j=[(\log k)^{1/4}]+1$, noting that when $1\le J<j$, one has

$$\Omega_J=\sum_{l>\widetilde{r}_J}k_l=\sum_{k_l\le J}k_l\le\sum_{1\le l\le J}l=\tfrac12J(J+1)<(\log k)^{1/2}.$$

Then on writing ϕ_J for $\phi^*(j,s,J)$, and noting that $\widetilde{r}_J\le t$ for each J, we deduce from (6.3) that

$$\phi_{J-1}\le(2k)^{-1}+\tfrac12(1-\delta')\phi_J, \tag{6.16}$$

where

$$\delta'=(\Delta_{s-t}-(\log k)^{1/2})/(kt). \tag{6.17}$$

Thus, starting with $\phi_j=1/k$ and following the obvious downwards induction using (6.16), we obtain

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

Since $\Delta_{s-t}\ge(\log k)^2$, we deduce from (6.17) that

$$\delta'\ge\delta_{s-t}\left(1-(\log k)^{-3/2}\right),$$

and therefore

$$\phi_1\le\frac{1+2^{1-j}\delta'}{k(1+\delta')}<\frac{1+2\delta_{s-t}(\log k)^{-3/2}}{k(1+\delta_{s-t})}. \tag{6.18}$$

Write $\theta=\phi_1$. Then Theorem 6.1 implies that the exponent Δ_s is admissible, where by (6.7),

$$\Delta_s=\Delta_{s-t}(1-\theta)+t(k\theta-1).$$

On writing δ_s for $\Delta_s/(tk)$, it follows from (6.18) that

$$\begin{aligned}\delta_s&=\delta_{s-t}-1/k+(1-\delta_{s-t})\theta\\&\le\delta_{s-t}-1/k+\frac{(1-\delta_{s-t})}{k(1+\delta_{s-t})}\left(1+2\delta_{s-t}(\log k)^{-3/2}\right)\\&=\delta_{s-t}\left(1-\frac{2-\omega}{k(1+\delta_{s-t})}\right),\end{aligned}$$

where $\omega = 2(1-\delta_{s-t})(\log k)^{-3/2}$. Consequently

$$\begin{aligned}\delta_s + \log \delta_s &\le \delta_{s-t}\left(1 - \frac{2-\omega}{k(1+\delta_{s-t})}\right) + \log \delta_{s-t} + \log\left(1 - \frac{2-\omega}{k(1+\delta_{s-t})}\right)\\ &\le \delta_{s-t} + \log \delta_{s-t} - \frac{2-\omega}{k(1+\delta_{s-t})}\delta_{s-t} - \frac{2-\omega}{k(1+\delta_{s-t})}\\ &= \delta_{s-t} + \log \delta_{s-t} - (2-\omega)/k,\end{aligned}$$

whence the desired conclusion (6.15) follows immediately.

Theorems 2 and 3 may now be established by applying Lemma 6.2 inductively.

The proof of Theorem 2. Again we suppose that k_1 is sufficiently large, and write $k = k_1$. When $s \in \mathbb{N}$, define δ_s to be the unique positive solution of the equation

$$\delta_s + \log \delta_s = 1 - \frac{2s}{tk} + \frac{2s}{tk(\log k)^{3/2}}. \tag{6.19}$$

We claim that the exponent $\Delta_s = tk\delta_s$ is admissible whenever $1 \le s \le t+1$ or $\delta_{s-t} > (\log k)^2/(tk)$.

Observe first that when r is an integer with $1 \le r \le t+1$, then by applying Hölder's inequality in combination with Theorem 1 of [25], it follows that the exponent $\Delta_r = \left(\sum_{i=1}^t k_i\right) - r$ is admissible. Thus on writing $\gamma_r = \Delta_r/(kt)$, we deduce from (6.19) that

$$\gamma_r + \log \gamma_r \le 1 - r/(tk) + \log(1 - r/(tk)) < 1 - 2r/(tk) < \delta_r + \log \delta_r,$$

whence our claim follows in the case $s = r$.

Next we observe that whenever $1 \le s \le s_0$, with $s_0 = \frac{1}{2}tk\,(\log(tk) - 2\log\log k)$, then by (6.19) one has $0 \le \delta_s < 1$ and

$$\delta_s + \log \delta_s \ge 1 - \frac{2s_0}{tk} + \frac{2s_0}{tk(\log k)^{3/2}} > 1 - \log(tk) + 2\log\log k,$$

whence $\delta_s > (\log k)^2/(tk)$. It therefore follows from Lemma 6.2 that whenever $1 \le s \le s_0$ and the exponent $\Delta_{s-t} = tk\delta_{s-t}$ is admissible, then so is the exponent $\Delta_s = tk\gamma_s$, where γ_s is the unique positive solution of the equation

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

But by applying (6.19), one obtains

$$\gamma_s + \log \gamma_s = 1 - \frac{2s}{tk} + \frac{2s}{tk(\log k)^{3/2}} = \delta_s + \log \delta_s,$$

so that $\gamma_s = \delta_s$. Our assertion therefore follows by induction, using the conclusion of the previous paragraph as the basis of the induction.

The conclusion of Theorem 2 follows immediately from the above assertion when $1 \le s \le s_0$. We next consider an integer s with $s > s_0$, and denote by U the largest integer with $U < s_0$ such that $U \equiv s \pmod{t}$. Then $U \ge s_0 - t$, and it follows from the assertion established above that the exponent Δ_U is admissible, where

$$\Delta_U = tke^{2-2U/(tk)} < e^3(\log k)^2.$$

Thus it follows from Theorem 5.2 that the exponent Δ_s is admissible, where

$$\Delta_s = \Delta_U(1-1/k)^{(s-U)/t} < e^3(\log k)^2(1-1/k)^{(s-s_0)/t}.$$

This completes the proof of Theorem 2.

The proof of Theorem 3. We start by making an elementary observation concerning admissible exponents for associated systems. When $1 \le r \le t$, write $\boldsymbol{\kappa}_r$ for the r-tuple $(k_1, \dots, k_r)$, and denote by $S_{u,r}$ the number of solutions of the system

$$\sum_{i=1}^{u} (x_i^{k_j} - y_i^{k_j}) = 0 \quad (1 \le j \le r), \tag{6.20}$$

with $x_i, y_i \in \mathcal{A}(P, R)$ $(1 \le i \le u)$. In view of the trivial inequality $S_{u,r+1} \le S_{u,r}$, whenever $\lambda_{u,\boldsymbol{\kappa}_r}$ is a permissible exponent, one has $S_{u,r+1} \ll P^{\lambda_{u,\boldsymbol{\kappa}_r}+\varepsilon}$, whence the exponent $\lambda_{u,\boldsymbol{\kappa}_{r+1}}$ is permissible with $\lambda_{u,\boldsymbol{\kappa}_{r+1}} = \lambda_{u,\boldsymbol{\kappa}_r}$. It follows that whenever the exponent $\Delta_{u,\boldsymbol{\kappa}_r}$ is admissible, then so is $\Delta_{u,\boldsymbol{\kappa}_{r+1}}$, where

$$\Delta_{u,\boldsymbol{\kappa}_{r+1}} = \Delta_{u,\boldsymbol{\kappa}_r} + k_{r+1}. \tag{6.21}$$

Suppose that $k_1, \dots, k_t$ satisfy (1.2) and that k_1 is sufficiently large. We adopt the convention that $\boldsymbol{\kappa}_0 = (k_1)$, and put $v_0 = 0$. Also, when $1 \le j \le t-1$ we put

$$v_j = \left[\tfrac{1}{2} k_1 \log\left(e^3 k_j/k_{j+1}\right) + 1\right], \tag{6.22}$$

and write

$$U_j = \sum_{i=0}^{j} i v_i. \tag{6.23}$$

We claim that when $0 \le j \le t-1$, the exponent $\Delta_{U_j,\boldsymbol{\kappa}_j}$ is admissible, where

$$\Delta_{U_j,\boldsymbol{\kappa}_j} = \max\left\{k_{j+1}, (\log k_1)^2\right\}. \tag{6.24}$$

We prove this assertion by induction, noting that it is trivial when $j = 0$.

Suppose that $1 \le J \le t-1$, and that our assertion holds for $j = J-1$. Then the exponent $\Delta_{U_{J-1},\boldsymbol{\kappa}_{J-1}} = \max\left\{k_J, (\log k_1)^2\right\}$ is admissible, whence by our opening observation, the exponent $\Delta_{U_{J-1},\boldsymbol{\kappa}_J} = 2\max\left\{k_J, (\log k_1)^2\right\}$ is also admissible. On applying Lemma 6.2 and recalling (6.23), we deduce that the exponent $\Delta_{U_J,\boldsymbol{\kappa}_J} = J k_1 \delta_{U_J}$ is admissible, where

$$\delta_{U_J} + \log \delta_{U_J} = \frac{2\max\left\{k_J, (\log k_1)^2\right\}}{J k_1} + \log\left(\frac{2\max\left\{k_J, (\log k_1)^2\right\}}{J k_1}\right) - \mathcal{E}_J,$$

and

$$\mathcal{E}_J = \frac{2 v_J}{k_1}\left(1 - (\log k_1)^{-3/2}\right).$$

Moreover, by (6.22) one has $\mathcal{E}_J > \log(2e^{2/J} k_J/k_{J+1})$, whence

$$\log \delta_{U_J} < \log\left(\max\left\{k_{J+1}, (\log k_1)^2\right\}/(J k_1)\right).$$

Thus the exponent $\Delta_{U_J,\boldsymbol{\kappa}_J} = \max\left\{k_{J+1}, (\log k_1)^2\right\}$ is admissible, and our assertion follows for $j = J$, completing the induction.

Adopt the convention that $k_{t+1} = (\log k_1)^2$, and write

$$v_t = \left[\tfrac{1}{2} k_1 \log\left(e^3 k_t\right) + 1\right] \quad \text{and} \quad U = \sum_{i=0}^{t} i v_i.$$

We observe that the argument of the previous paragraph leads to the conclusion that the exponent $\Delta_{U,\boldsymbol{\kappa}_t} = (\log k_1)^2$ is admissible. Furthermore Theorem 5.2 implies that whenever $s_1 \ge U$, and s is an integer exceeding s_1 with $s \equiv s_1 \pmod{t}$, then the exponent $\Delta_{s,\mathbf{k}} = (\log k_1)^2 (1 - 1/k_1)^{(s-s_1)/t}$ is admissible. The proof of Theorem 3 is therefore completed by noting that

$$\begin{aligned} U &\le \sum_{j=1}^{t-1} \left(\tfrac{1}{2} j k_1 \log(e^3 k_j/k_{j+1}) + j\right) + \tfrac{1}{2} t k_1 \log(e^3 k_t) + t \\ &< \tfrac{1}{2} k_1 \log(k_1 \dots k_t) + \tfrac{3}{4} t(t+1) k_1 + \tfrac{1}{2} t(t+1), \\ &< \tfrac{1}{2} k_1 \left(\log(k_1 \dots k_t) + 3t^2\right). \end{aligned}$$

Part II: Several applications

7. Estimates for smooth Weyl sums

The estimates for smooth Weyl sums provided by Theorems 4 and 5 are fairly immediate consequences of Lemma 7.3, which we prove using an argument closely related to that used to establish [28], Lemma 4.1 (itself similar to the proof of Vaughan [19], Theorem 1.8). In order to prove Lemma 7.3 we will require two preliminary lemmata. Recall the notation defined at the end of §2, and define the set $\mathcal{C}_q(Q)$ by

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

Lemma 7.1. *Let L be a positive real number and r be a positive integer with $\log r \ll \log L$. Then for each $\varepsilon > 0$, one has card$(\mathcal{C}_r(L)) \ll L^\varepsilon$.*

Proof. This is [28], Lemma 2.2.

Before stating the second preliminary lemma, we record some notation. Write $\psi(x; \boldsymbol{\alpha}) = \sum_{i=1}^t \alpha_i x^{k_i}$, and define the exponential sum $h_{r,v}(\boldsymbol{\alpha}; L, R; \theta)$ by

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

Also, when π is a prime number, define the set of modified smooth numbers $\mathcal{B}(M, \pi, R)$ by

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

Lemma 7.2. *Let $\boldsymbol{\alpha} \in \mathbb{R}^t$ and $r \in \mathbb{N}$, and suppose that Q, M and R satisfy $2 \le R \le M < Q$. Then*

$$\sum_{\substack{x \in \mathcal{A}(Q,R) \\ (x,r)=1}} e(\psi(x; \boldsymbol{\alpha})) \ll R \log Q \max_{\substack{\pi \le R \\ \pi \text{ prime}}} \sup_{\theta \in [0,1)} \sum_{\substack{v \in \mathcal{B}(M,\pi,R) \\ (v,r)=1}} |h_{r,v}(\boldsymbol{\alpha}; Q/M, \pi; \theta)| + M.$$

Proof. One may employ the argument of the proof of Vaughan [19], Lemma 10.1 precisely as in the treatment of Wooley [28], Lemma 2.3.

Lemma 7.3. *Let $k_1, \dots, k_t$ be integers satisfying* (1.2). *Suppose that λ is a real number with $0 < \lambda \le \frac{1}{2}$, and write $M = P^\lambda$. Let j be an integer with $1 \le j \le t$, and let $\boldsymbol{\alpha} \in \mathbb{R}^t$. Suppose that $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $(a, q) = 1$, $|q\alpha_j - a| \le \frac{1}{2}(MR)^{-k_j}$, $q \le 2(MR)^{k_j}$, and either $|q\alpha_j - a| > MP^{-k_j}$ or $q > MR$. Then whenever s is a natural number with $2s > k_1$, and the exponent Δ_s is admissible,*

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

Proof. We start by manipulating $f_{\mathbf{k}}(\boldsymbol{\alpha}; P, R)$ into a form suitable for our subsequent application of the large sieve inequality. On observing that

$$f_{\mathbf{k}}(\boldsymbol{\alpha}; P, R) = \sum_{d \in \mathcal{C}_q(P) \cap \mathcal{A}(P,R)} \sum_{\substack{x \in \mathcal{A}(P/d,R) \\ (x,q)=1}} e(\psi(xd; \boldsymbol{\alpha})),$$

we deduce from Lemma 7.1 that

$$f_{\mathbf{k}}(\boldsymbol{\alpha}; P, R) \ll P^\varepsilon \max_{d \in \mathcal{C}_q(M/R)} \left| \sum_{\substack{x \in \mathcal{A}(P/d,R) \\ (x,q)=1}} e(\psi(xd; \boldsymbol{\alpha})) \right| + (PR/M)^{1+\varepsilon}.$$

Then it follows from Lemma 7.2 that there exists $d \in \mathcal{C}_q(M/R)$, $\theta \in [0, 1)$, and a prime π with $\pi \le R$ such that

$$f_{\mathbf{k}}(\boldsymbol{\alpha}; P, R) \ll P^{1+\varepsilon} M^{-1} + P^\varepsilon g(\boldsymbol{\alpha}; d, \pi, \theta), \tag{7.1}$$

where

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

Denote by $J(q,d,h)$ the number of solutions of the congruence $(xd)^{k_j}\equiv h\pmod q$ with $1\le x\le q$, and note that when $(h,q)|d^{k_j}$ one has $J(q,d,h)\ll q^{\varepsilon}d^{k_j}$. Then there exists an integer L with $L\ll q^{\varepsilon}d^{k_j}$ such that the integers v with $M/d<v\le MR/d$ and $(v,q)=1$ can be divided into L classes $\mathcal{V}_1,\dots,\mathcal{V}_L$, with the property that for any two distinct elements v_1,v_2 in a given class $\mathcal{V}_r$, we have $(v_1d)^{k_j}\equiv(v_2d)^{k_j}\pmod q$ if and only if $v_1\equiv v_2\pmod q$. Consequently, on writing $c_{\mathbf{y}}$ for the number of solutions of the diophantine system

$$\sum_{i=1}^{s}u_i^{k_j}\equiv y_j\quad(1\le j\le t),$$

with $u_i\in\mathcal{A}(P/M,\pi)$ $(1\le i\le s)$, we may apply Hölder's inequality to obtain

$$g(\boldsymbol{\alpha};d,\pi,\theta)^{2s}\ll P^{\varepsilon}d^{k_j}(MR/d)^{2s-1}\max_{1\le r\le L}\sum_{v\in\mathcal{V}_r}\left|\sum_{\mathbf{y}}b_{\mathbf{y}}e(\psi(vd;\boldsymbol{\alpha}\mathbf{y}))\right|^2,$$

where $|b_{\mathbf{y}}|\le c_{\mathbf{y}}$, the summation is over $\mathbf{y}$ with $1\le y_i\le s(P/M)^{k_i}$ $(1\le i\le t)$, and we have written $\boldsymbol{\alpha}\mathbf{y}$ for the t-tuple $(\alpha_1y_1,\dots,\alpha_ty_t)$. Then by Cauchy's inequality,

$$g(\boldsymbol{\alpha};d,\pi,\theta)^{2s}\ll P^{\varepsilon}d^{k_j}(MR/d)^{2s-1}\left(\prod_{l\ne j}s(P/M)^{k_l}\right)\mathcal{W}_j,\tag{7.2}$$

where

$$\mathcal{W}_j={\sum_{\mathbf{y}}}^{*}\max_{1\le r\le L}\sum_{v\in\mathcal{V}_r}\left|\sum_{1\le y_j\le s(P/M)^{k_j}}b_{\mathbf{y}}e(\alpha_j(vd)^{k_j}y_j)\right|^2,$$

and the first summation is over y_l satisfying $1\le y_l\le s(P/M)^{k_l}$ $(1\le l\le t,\ l\ne j)$.

Next we show that the $\alpha_j(vd)^{k_j}$ are somewhat widely spaced apart modulo 1. It is this property which is fundamental to the strength of the estimates stemming from our later application of the large sieve inequality. We start by observing that if $v_1,v_2\in\mathcal{V}_r$ and $v_1\not\equiv v_2\pmod q$, then one has $(v_1d)^{k_j}\not\equiv(v_2d)^{k_j}\pmod q$, and hence the inequality $|q\alpha_j-a|\le\frac12(MR)^{-k_j}$ implies that

$$\|\alpha_j\left((v_1d)^{k_j}-(v_2d)^{k_j}\right)\|\ge\left\|\frac{a}{q}\left((v_1d)^{k_j}-(v_2d)^{k_j}\right)\right\|-\tfrac12q^{-1}\ge\tfrac12q^{-1}.\tag{7.3}$$

We divide into cases.

(i) Suppose that $q>MR/d$. Then the elements of $\mathcal{V}_r$ are distinct modulo q, and so it follows from (7.3) that for $v\in\mathcal{V}_r$, the $\alpha_j(vd)^{k_j}$ are spaced at least $\frac12q^{-1}$ apart modulo 1.

(ii) Suppose that $q\le MR/d$. Then by hypothesis we have $|q\alpha_j-a|>MP^{-k_j}$. Given any two elements v_1,v_2 of $\mathcal{V}_r$ with $v_1\not\equiv v_2\pmod q$, we may conclude as in case (i) that $\alpha_j(v_1d)^{k_j}$ and $\alpha_j(v_2d)^{k_j}$ are spaced at least $\frac12q^{-1}$ apart modulo 1. Meanwhile, when $v_1\equiv v_2\pmod q$ but $v_1\ne v_2$, one has

$$\|\alpha_j\left((v_1d)^{k_j}-(v_2d)^{k_j}\right)\|=\|(\alpha_j-a/q)(v_1^{k_j}-v_2^{k_j})d^{k_j}\|=|\alpha_j-a/q|\cdot|v_1^{k_j}-v_2^{k_j}|d^{k_j}.$$

Then since v_1-v_2 is a non-zero multiple of q, and $v_id>M$ $(i=1,2)$, we obtain

$$\|\alpha_j\left((v_1d)^{k_j}-(v_2d)^{k_j}\right)\|>|\alpha_j-a/q|\,M^{k_j-1}q>(P/M)^{-k_j}.$$

We therefore deduce that in this case the $\alpha_j(vd)^{k_j}$ $(v\in\mathcal{V}_r)$ are spaced at least $\frac12\min\left\{q^{-1},(P/M)^{-k_j}\right\}$ apart modulo 1.

Thus in either case the large sieve inequality (see, for example, Vaughan [18], Lemma 5.3) implies that

$$\sum_{\mathbf{y}}^{*} \max_{1\le r\le L} \sum_{v\in\mathcal{V}_r} \Bigg| \sum_{1\le y_j\le s(P/M)^{k_j}} b_{\mathbf{y}} e(\alpha_j (vd)^{k_j} y_j) \Bigg|^2 \ll \big(q+(P/M)^{k_j}\big) \sum_{\mathbf{y}} |b_{\mathbf{y}}|^2$$
$$\ll (q+(P/M)^{k_j}) S_s(P/M,R).$$

Then since $2s \ge k_j+1$ and Δ_s is an admissible exponent, it follows from (7.1) and (7.2) that

$$f_{\mathbf{k}}(\boldsymbol{\alpha};P,R)^{2s} \ll P^{\varepsilon} M^{2s-1} (P/M)^{2s+\Delta_s} (1+q(P/M)^{-k_j}) + (P/M)^{2s+\varepsilon}.$$

The lemma follows on noting that the second term on the right hand side of the last inequality is smaller than the first.

The proof of Theorem 4. Recall the notation of the statement of Theorem 4. We suppose that λ is a real number with $0<\lambda\le 1/t$, and that $\boldsymbol{\alpha}\in\mathfrak{m}_{t\lambda}$. For convenience we write $M=P^{\lambda}$. For each i with $1\le i\le t$, it follows from Dirichlet's Theorem that there exist $b_i\in\mathbb{Z}$ and $q_i\in\mathbb{N}$ with $(b_i,q_i)=1$, $q_i\le 2(MR)^{k_i}$ and $|q_i\alpha_i-b_i|\le \frac12 (MR)^{-k_i}$. If for some j with $1\le j\le t$ we have $|q_j\alpha_j-b_j|>MP^{-k_j}$ or $q_j>MR$, then we may apply Lemma 7.3 to obtain

$$f_{\mathbf{k}}(\boldsymbol{\alpha};P,R) \ll P^{1+\varepsilon}\left(M^{-1}(P/M)^{\Delta_s}\right)^{1/(2s)} \ll P^{1-\sigma(\mathbf{k};\lambda)+\varepsilon},$$

and we are done. Otherwise, for every i with $1\le i\le t$ we have $q_i\le MR$ and $|q_i\alpha_i-b_i|\le MP^{-k_i}$. We put $q=[q_1,\dots,q_t]$ and $a_i=b_iq/q_i$ $(1\le i\le t)$. Then $(a_1,\dots,a_t,q)=1$, $q\le (MR)^t\le P^{t\lambda}R^t$, and

$$|q\alpha_i-a_i|\le MP^{-k_i}(MR)^{t-1}\le P^{t\lambda-k_i}R^{t-1}\quad (1\le i\le t).$$

Thus $\boldsymbol{\alpha}\notin\mathfrak{m}_{t\lambda}$, and we derive a contradiction. The theorem follows immediately.

The argument used to prove Theorem 5 is essentially the same as that used in the proof of Theorem 4, save that the lower dimension of the problem leads to simplifications.

The proof of Theorem 5. Recall the notation of the statement of Theorem 5. We suppose that λ is a real number with $0<\lambda\le\frac12$, and that $\alpha\in\mathfrak{n}_{\lambda}$. For convenience we write $M=P^{\lambda}$. By Dirichlet's Theorem there exist $b\in\mathbb{Z}$ and $r\in\mathbb{N}$ with $(b,r)=1$, $r\le 2(MR)^{k_1}$ and $|ra_1\alpha-b|\le\frac12(MR)^{-k_1}$. If $|ra_1\alpha-b|>MP^{-k_1}$ or $r>MR$, then we may apply Lemma 7.3 to obtain

$$g_{\mathbf{k}}(\alpha;P,R)=f_{\mathbf{k}}(\mathbf{a}\alpha;P,R)\ll P^{1+\varepsilon}\left(M^{-1}(P/M)^{\Delta_s}\right)^{1/(2s)}\ll P^{1-\sigma(\mathbf{k};\lambda)+\varepsilon},$$

and we are done. Otherwise we have $r\le MR$ and $|ra_1\alpha-b|\le MP^{-k_1}$. We put

$$q=\frac{r|a_1|}{(b,a_1)}\quad\text{and}\quad a=\frac{b|a_1|}{(b,a_1)a_1}.$$

Then $(a,q)=1$, $q\le |a_1|P^{\lambda}R$ and $|q\alpha-a|\le P^{\lambda-k_1}$. Thus $\alpha\notin\mathfrak{n}_{\lambda}$, a contradiction which completes the proof of the theorem.

The proof of Theorem 6. We establish the two parts of Theorem 6 through separate though similar arguments, beginning with the first part. Recall the notation of the statement of Theorem 6, together with the definition of $\psi(x;\boldsymbol{\alpha})$ given at the beginning of this section. Write $\tau=\tau(\mathbf{k})$, and let ε and ϕ be real numbers with $0<\varepsilon<\phi<\tau$. Let P be a sufficiently large real number, and write $H_1=P^{\tau-\phi}$. Define $T_1(\boldsymbol{\alpha})$ by

$$T_1(\boldsymbol{\alpha})=\sum_{1\le h\le H_1}|f_{\mathbf{k}}(h\boldsymbol{\alpha};P,R)|.$$

Then it follows from Lemma 5 of [5] that the estimate $\min_{1\le n\le P}\|\psi(n;\boldsymbol{\alpha})\|<H_1^{-1}$ is implied by the bound $T_1(\boldsymbol{\alpha})=o(P)$. We exploit this observation by applying Theorem 4 with $\lambda=(1-\tau)/t$. Suppose

first that there exist $h,\mathbf{b},q$ with $1\le h\le H_1$, $\mathbf{b}\in\mathbb{Z}^t$, $q\in\mathbb{N}$, $(b_1,\dots,b_t,q)=1$, $q\le P^{1-\tau}R^t$ and $|qh\alpha_i-b_i|\le P^{1-\tau-k_i}R^{t-1}$ $(1\le i\le t)$. Then for each i we have

$$\|\alpha_i(qh)^{k_i}\|\le|(qh)^{k_i-1}(qh\alpha_i-b_i)|<(H_1P^{1-\tau}R^t)^{k_i-1}P^{1-\tau-k_i}R^{t-1}.$$

Then since $(\tau-\phi)+1-\tau+t\eta<1$, we obtain $\|\alpha_i(qh)^{k_i}\|<t^{-1}H_1^{-1}$, whence on noting that $qH_1\le P^{1-\phi}R^t$, we deduce that

$$\min_{1\le n\le P}\|\psi(n;\boldsymbol{\alpha})\|\le\|\psi(qh;\boldsymbol{\alpha})\|\le\sum_{i=1}^t\|\alpha_i(qh)^{k_i}\|<H_1^{-1}.$$

Meanwhile, if for each $h,\mathbf{b},q$ with $1\le h\le H_1$, $\mathbf{b}\in\mathbb{Z}^t$, $q\in\mathbb{N}$, $(b_1,\dots,b_t,q)=1$ and $|qh\alpha_i-b_i|\le P^{1-\tau-k_i}R^{t-1}$ $(1\le i\le t)$, one has $q>P^{1-\tau}R^t$, then Theorem 4 implies that

$$\max_{1\le h\le H_1}|f_{\mathbf{k}}(h\boldsymbol{\alpha};P,R)|\ll P^{1-\sigma(\mathbf{k};\lambda)+\varepsilon}.$$

Moreover, a simple calculation reveals that $\sigma(\mathbf{k};\lambda)\ge\tau$, so that

$$T_1(\boldsymbol{\alpha})\ll H_1P^{1-\tau+\varepsilon}\ll P^{1+\varepsilon-\phi}=o(P).$$

Thus our earlier observation once again shows that $\min_{1\le n\le P}\|\psi(n;\boldsymbol{\alpha})\|<H_1^{-1}$, and the first part of Theorem 6 follows immediately.

For the second part of Theorem 6, we put $\boldsymbol{\alpha}=\alpha\mathbf{a}$ for some rational t-tuple $\mathbf{a}$, and write $\sigma_2=\sigma(\mathbf{k};1/2)$. Let ε and ϕ be real numbers with $0<\varepsilon<\phi<\sigma_2$, let P be a sufficiently large real number, and write $H_2=P^{\sigma_2-\phi}$. Define $T_2(\alpha)$ by

$$T_2(\alpha)=\sum_{1\le h\le H_2}|g_{\mathbf{k}}(h\alpha;P,R)|.$$

Once again we deduce from Lemma 5 of [5] that in order to establish the bound $\min_{1\le n\le P}\|\psi(n;\boldsymbol{\alpha})\|<H_2^{-1}$, it suffices to show that $T_2(\alpha)=o(P)$. We apply Theorem 5 with $\lambda=\frac{1}{2}$. Let d be the least positive integer such that $d\mathbf{a}\in\mathbb{Z}^t$, and write $\beta=\alpha/d$. We suppose that there exist h, b, q with $1\le h\le dH_2$, $b\in\mathbb{Z}$, $q\in\mathbb{N}$, $(b,q)=1$, $q\le|a_1d|P^\lambda R$ and $|qh\beta-b|\le P^{\lambda-k_1}$. Then

$$\|\beta(a_id)(qh)^{k_i}\|\le|(qh)^{k_i-1}(qh\beta-b)(a_id)|<|a_id|\left(|a_1d^2|H_2P^\lambda R\right)^{k_i-1}P^{\lambda-k_1}.$$

Then $\|\beta(a_id)(qh)^{k_i}\|<t^{-1}H_2^{-1}$, whence

$$\min_{1\le n\le P}\|\psi(n;\boldsymbol{\alpha})\|\le\|\psi(qh;\boldsymbol{\alpha})\|\le\sum_{i=1}^t\|\beta(a_id)(qh)^{k_i}\|<H_2^{-1}.$$

Meanwhile, if for each h, b, q with $1\le h\le dH_2$, $b\in\mathbb{Z}$, $q\in\mathbb{N}$, $(b,q)=1$ and $|qh\beta-b|\le P^{\lambda-k_1}$, one has $q>|a_1d|P^\lambda R$, then Theorem 5 implies that

$$\max_{1\le h\le H_2}|g_{\mathbf{k}}(h\alpha;P,R)|\le\max_{1\le h\le dH_2}|g_{\mathbf{k}}(h\beta;P,R)|\ll P^{1-\sigma_2+\varepsilon},$$

whence

$$T_2(\alpha)\ll H_2P^{1-\sigma_2+\varepsilon}\ll P^{1+\varepsilon-\phi}=o(P).$$

Then the second part of Theorem 6 follows in this case also, and the proof of the theorem is complete.

The consequences of Theorems 4, 5 and 6 recorded in Theorem 7 are almost immediate from Theorems 2 and 3. We will therefore be brief in our proof of Theorem 7.

The proof of Theorem 7. We start by proving part (i) of the theorem. We put

$$u_0=\left[\tfrac{1}{2}tk_1(\log k_1+4\log\log k_1+3\log t+6)\right]+1,$$

and

$$u_1 = \left[\tfrac{1}{2}k_1(\log(k_1\dots k_t)+3t^2)\right] + t\,[3k_1\log\log k_1 + k_1\log t + 1]\,.$$

On recalling the notation of Theorems 2 and 3, we have

$$u_0 - s_0 > tk_1\,(3\log\log k_1 + \log t + 3) \quad \text{and} \quad u_1 - s_1 > tk_1(3\log\log k_1 + \log t),$$

and hence the latter theorems imply that the exponents Δ_{u_0} and Δ_{u_1} are each admissible, where

$$\Delta_{u_0} = e^3(\log k_1)^2\exp\left((s_0-u_0)/(tk_1)\right) < 1/(t\log k_1), \tag{7.4}$$

and

$$\Delta_{u_1} = (\log k_1)^2\exp\left((s_1-u_1)/(tk_1)\right) < 1/(t\log k_1). \tag{7.5}$$

On applying Theorem 4, therefore, we deduce that

$$\sup_{\boldsymbol{\alpha}\in\mathfrak{m}_1} |f_{\mathbf{k}}(\boldsymbol{\alpha};P,R)| \ll P^{1-\sigma(\mathbf{k};1/t)+\varepsilon},$$

where

$$\sigma(\mathbf{k};1/t) \ge \max_{i=0,1}\frac{1-(t-1)\Delta_{u_i}}{2tu_i}. \tag{7.6}$$

On substituting (7.4) and (7.5) into (7.6), we obtain

$$\sigma(\mathbf{k};1/t)^{-1} < 2t\left(1-1/(\log k_1)\right)^{-1}\min\{u_0,u_1\}\,,$$

whence the first estimate of part (i) follows immediately.

Next we put

$$v_0 = \left[\tfrac{1}{2}tk_1(\log k_1 + 4\log\log k_1 + \log t + 6)\right] + 1,$$

and

$$v_1 = \left[\tfrac{1}{2}k_1\left(\log(k_1\dots k_t)+3t^2\right)\right] + t\,[3k_1\log\log k_1 + 1]\,.$$

Following the argument of the first paragraph, we deduce that the exponents Δ_{v_0} and Δ_{v_1} are admissible, where

$$\max\{\Delta_{v_0},\Delta_{v_1}\} < 1/(\log k_1). \tag{7.7}$$

Thus Theorem 5 leads to the estimate

$$\sup_{\alpha\in\mathfrak{n}_{1/2}} |g_{\mathbf{k}}(\alpha;P,R)| \ll P^{1-\sigma(\mathbf{k};1/2)+\varepsilon},$$

where

$$\sigma(\mathbf{k};1/2) \ge \max_{i=0,1}\frac{1-\Delta_{v_i}}{4v_i}. \tag{7.8}$$

On substituting (7.7) into (7.8) we establish the second estimate of part (i).

In order to complete the proof of the theorem, we have merely to note that part (ii) follows from Theorem 6 in combination with the estimates for $\sigma(\mathbf{k};1/t)$ and $\sigma(\mathbf{k};1/2)$ provided by (7.6) and (7.8), since $\tau(\mathbf{k}) = (1+o(1))\sigma(\mathbf{k};1/t)$.

8. Upper bounds for large moduli

The unlocalised fractional parts estimates provided by Theorem 8 are simple consequences of an exponential sum estimate related to that established in [28], Lemma 3.1 by using a method similar to one of Heath-Brown [7], §5. We prove this estimate in the following theorem.

Theorem 8.1. *Let $k_1,\dots,k_t$ be integers satisfying* (1.2)*, and let H be a real number with $H\ge 1$. Suppose that λ is a real number with $\frac{1}{2}\le\lambda<1$, and write $M=P^\lambda$. Let $\boldsymbol{\alpha}\in\mathbb{R}^t$, $\mathbf{b}\in\mathbb{Z}^t$, and suppose that there exist $\mathbf{a}\in\mathbb{Z}^t$ and $\mathbf{q}\in\mathbb{N}^t$ satisfying the conditions $(a_i,q_i)=1$ and $|\alpha_i-a_i/q_i|\le q_i^{-2}$ $(1\le i\le t)$. Then whenever m and w are natural numbers, and Δ_m and Δ_w are admissible exponents,*

$$\sum_{1\le h\le H}|f_{\mathbf{k}}(h\mathbf{b}\boldsymbol{\alpha};P,R)|\ll_{\mathbf{b}}(HP)^{1+\varepsilon}\left(M^{\Delta_w}(P/M)^{\Delta_m}\Theta_{\mathbf{k}}(P;\mathbf{q};\lambda)\right)^{1/(2mw)}+HM,$$

where $\mathbf{b}\boldsymbol{\alpha}$ denotes $(b_1\alpha_1,\dots,b_t\alpha_t)$,

$$\Theta_{\mathbf{k}}(P;\mathbf{q};\lambda)=H^{-1}\min_{\substack{\mathcal{J}\subseteq\{1,\dots,t\}\\ \mathcal{J}\neq\emptyset}}\prod_{i\in\mathcal{J}}\theta_i(P;\lambda),$$

and

$$\theta_i(P;\lambda)=Hq_i^\varepsilon\left(q_i^{-1}+(P/M)^{-k_i}+q_iH^{-1}P^{-k_i}\right).$$

Proof. We first consider a fixed integer h with $1\le h\le H$, and rewrite the exponential sum $f_{\mathbf{k}}(h\mathbf{b}\boldsymbol{\alpha};P,R)$ in a form suitable for our later applications of Hölder's inequality. For convenience we write $\boldsymbol{\beta}=h\mathbf{b}\boldsymbol{\alpha}$. By applying Lemma 7.2 with $r=1$, we deduce that there exists a prime π with $\pi\le R$, and $\theta\in[0,1)$ such that

$$f_{\mathbf{k}}(\boldsymbol{\beta};P,R)\ll P^\varepsilon Rg(\boldsymbol{\beta};\theta)+M,\tag{8.1}$$

where

$$g(\boldsymbol{\beta};\theta)=\sum_{v\in\mathcal{A}(MR,R)}|h_{1,v}(\boldsymbol{\beta};P/M,\pi;\theta)|.$$

Define the complex numbers of unit modulus, $\varepsilon(v,\theta)$, by

$$|h_{1,v}(\boldsymbol{\beta};P/M,\pi;\theta)|^m=\varepsilon(v,\theta)h_{1,v}(\boldsymbol{\beta};P/M,\pi;\theta)^m.\tag{8.2}$$

Also, let $r(\mathbf{c};\theta)$ denote the number of solutions of the diophantine system

$$\sum_{i=1}^m u_i^{k_j}=c_j\quad(1\le j\le t),$$

with $u_i\in\mathcal{A}(P/M,\pi)$ $(1\le i\le m)$, in which each solution $\mathbf{u}$ is counted with weight $e(\theta(u_1+\dots+u_m))$. Thus

$$h_{1,v}(\boldsymbol{\beta};P/M,\pi;\theta)^m=\sum_{\mathbf{c}}r(\mathbf{c};\theta)e(\beta_1c_1v^{k_1}+\dots+\beta_tc_tv^{k_t}),$$

where the summation is over integral t-tuples $\mathbf{c}$ satisfying $1\le c_i\le m(P/M)^{k_i}$ $(1\le i\le t)$. An application of Hölder's inequality together with (8.2) therefore yields

$$g(\boldsymbol{\beta};\theta)^m\le(MR)^{m-1}\sum_{v\in\mathcal{A}(MR,R)}\varepsilon(v,\theta)\sum_{\mathbf{c}}r(\mathbf{c};\theta)e(\beta_1c_1v^{k_1}+\dots+\beta_tc_tv^{k_t}).$$

Next, on noting that $|r(\mathbf{c};\theta)|\le r(\mathbf{c};0)$, we find that a second application of Hölder's inequality reveals that

$$g(\boldsymbol{\beta};\theta)^{2mw}\le(MR)^{2w(m-1)}\left(\sum_{\mathbf{c}}r(\mathbf{c};0)\right)^{2w-2}\left(\sum_{\mathbf{c}}r(\mathbf{c};0)^2\right)J_w(\boldsymbol{\beta}),$$

where

$$J_w(\boldsymbol{\beta})=\sum_{\mathbf{c}}|\widetilde{g}(\boldsymbol{\beta};\mathbf{c},\theta)|^{2w},\tag{8.3}$$

and

$$\widetilde{g}(\boldsymbol{\beta};\mathbf{c},\theta)=\sum_{v\in\mathcal{A}(MR,R)}\varepsilon(v,\theta)e(\beta_1c_1v^{k_1}+\dots+\beta_tc_tv^{k_t}).$$

Consequently, on considering the underlying diophantine equations,

$$g(\boldsymbol{\beta};\theta)^{2mw}\le (PR)^{2mw}(P/M)^{-2m}(MR)^{-2w}S_m(P/M,R)J_w(\boldsymbol{\beta}).$$

Thus, on applying Hölder's inequality once more, and recalling (8.1), we obtain

$$\sum_{1\le h\le H}|f_{\mathbf{k}}(h\mathbf{b}\boldsymbol{\alpha};P,R)|\ll (HP)^{1+\varepsilon}\Psi^{1/(2mw)}+HM, \tag{8.4}$$

where

$$\Psi=H^{-1}(P/M)^{-2m}M^{-2w}S_m(P/M,R)\sum_{1\le h\le H}J_w(h\mathbf{b}\boldsymbol{\alpha}). \tag{8.5}$$

In order to complete our proof we must estimate Ψ. Let $n(\mathbf{d})$ denote the number of solutions of the system

$$\sum_{i=1}^w v_i^{k_j}-\sum_{i=w+1}^{2w}v_i^{k_j}=d_j\quad (1\le j\le t),$$

with $v_i\in\mathcal{A}(MR,R)$ $(1\le i\le 2w)$, in which each solution $\mathbf{v}$ is counted with weight $\prod_{i=1}^w\varepsilon(v_i,\theta)\overline{\varepsilon(v_{w+i},\theta)}$. Then

$$n(\mathbf{d})=\int_{\mathbb{T}^t}|\widetilde{g}(\boldsymbol{\beta};\mathbf{c},\theta)|^{2w}e(-\beta_1c_1d_1-\dots-\beta_tc_td_t)d\boldsymbol{\beta},$$

so that on recalling (8.3) we obtain

$$J_w(h\mathbf{b}\boldsymbol{\alpha})=\sum_{\mathbf{c}}\sum_{\mathbf{d}}n(\mathbf{d})e(h(\alpha_1b_1c_1d_1+\dots+\alpha_tb_tc_td_t)),$$

where the second summation is over $\mathbf{d}$ with $|d_i|\le w(MR)^{k_i}$ $(1\le i\le t)$. Thus, on interchanging the order of summation, and making the trivial estimate $|n(\mathbf{d})|\le n(\mathbf{0})\le S_w(MR,R)$, we deduce that

$$\begin{aligned}\sum_{1\le h\le H}J_w(h\mathbf{b}\boldsymbol{\alpha})&\le S_w(MR,R)\sum_{1\le h\le H}\sum_{\mathbf{d}}\prod_{i=1}^t\Biggl|\sum_{1\le c_i\le m(P/M)^{k_i}}e(\alpha_ihb_ic_id_i)\Biggr|\\ &\ll S_w(MR,R)(PR)^{k_1+k_2+\dots+k_t}\min_{\substack{\mathcal{J}\subseteq\{1,\dots,t\}\\ \mathcal{J}\ne\emptyset}}\prod_{i\in\mathcal{J}}\Upsilon_i(q_i),\end{aligned} \tag{8.6}$$

where

$$\Upsilon_i(q_i)=(PR)^{-k_i}\sum_{1\le h\le H}\sum_{|d_i|\le w(MR)^{k_i}}\min\left\{(P/M)^{k_i},\|hb_id_i\alpha_i\|^{-1}\right\}.$$

But [18], Lemma 2.2 implies that

$$\Upsilon_i(q_i)\ll_{\mathbf{b}}H^{1+\varepsilon}q_i^{\varepsilon}\left(q_i^{-1}+(P/M)^{-k_i}+H^{-1}(MR)^{-k_i}+q_iH^{-1}(PR)^{-k_i}\right),$$

whence by (8.5) and (8.6), on recalling that Δ_m and Δ_w are admissible exponents, we obtain

$$\Psi\ll (HP)^{\varepsilon}M^{\Delta_w}(P/M)^{\Delta_m}H^{-1}\min_{\substack{\mathcal{J}\subseteq\{1,\dots,t\}\\ \mathcal{J}\ne\emptyset}}\prod_{i\in\mathcal{J}}\theta_i(P;\lambda).$$

The lemma now follows immediately from (8.4).

The exponential sum estimate provided by Theorem 8.1 leads with little difficulty to a proof of Theorem 8 via standard fractional parts methods. Before proving the latter theorem, we remark that Theorem 8 is applicable in cases where the α_i are not all irrational. For suppose that the latter is the case. We partition the set of indices into two disjoint sets $\mathcal{Q}$ and $\mathcal{I}$ so that $\alpha_i\in\mathbb{Q}$ when $i\in\mathcal{Q}$, and $\alpha_j\notin\mathbb{Q}$ when $j\in\mathcal{I}$. Then there exist integers $a_i\in\mathbb{Z}$ and $q\in\mathbb{N}$ with $\alpha_i=a_i/q$ $(i\in\mathcal{Q})$, and moreover

$$\|\alpha_1(qn)^{k_1}+\dots+\alpha_t(qn)^{k_t}\|=\Biggl\|\sum_{i\in\mathcal{I}}\alpha_i(qn)^{k_i}\Biggr\|=\Biggl\|\sum_{i\in\mathcal{I}}\beta_in^{k_i}\Biggr\|,$$

where $\beta_i = \alpha_i q^{k_i}$ $(i \in \mathcal{I})$. Note that if $\mathcal{I} = \emptyset$, then $\|\alpha_1(qn)^{k_1} + \cdots + \alpha_t(qn)^{k_t}\| = 0$ for $n \in \mathbb{N}$, and we are done. In the remaining cases we deduce that whenever γ is a positive number for which, given any $\varepsilon > 0$, there are infinitely many solutions $n \in \mathbb{N}$ of the inequality $\left\|\sum_{i\in\mathcal{I}} \beta_i n^{k_i}\right\| \le n^{\varepsilon-\gamma}$, then given any $\varepsilon > 0$, there are infinitely many $n \in \mathbb{N}$ with $\left\|\sum_{i=1}^t \alpha_i n^{k_i}\right\| \le n^{\varepsilon-\gamma}$. Moreover the β_i are all irrational. It therefore follows from Theorem 8 that one may take

$$\gamma = \max_{s\in\mathbb{N}} \frac{\kappa_1 - 2\Delta_{s,\boldsymbol{\kappa}}}{4s^2},$$

where $\boldsymbol{\kappa} = (k_i)_{i\in\mathcal{I}}$, and $\kappa_1 = \max_{i\in\mathcal{I}} k_i$. Owing to the nature of our methods, in practice one will find that $\alpha(\mathbf{k}) \le \gamma$.

Proof of Theorem 8. We begin by disposing of the case in which the α_i are not all in rational ratio. Let a/q be a convergent to the continued fraction of α_1, so that $a \in \mathbb{Z}$, $q \in \mathbb{N}$, $(a,q) = 1$ and $|\alpha_1 - a/q| \le q^{-2}$. We define $\alpha(\mathbf{k})$ as in the statement of Theorem 8, with s the value of the parameter corresponding to the maximum. Let ϕ be any real number with $4\varepsilon < \phi < \alpha(\mathbf{k})$, and define the real numbers N and H by $N^{k_1-\phi} = q^2$ and $H = N^{\alpha(\mathbf{k})-\phi}$. We apply Theorem 8.1 with $\lambda = \frac{1}{2}$, $m = w = s$ and $b_i = 1$ $(1 \le i \le t)$, noting that we may take $\mathcal{J} = \{1\}$ in the implied minimum, and discard the remaining indices by taking any continued fraction convergents to the α_i with $i \ge 2$. We deduce that

$$\sum_{1\le h\le H} |f_{\mathbf{k}}(h\boldsymbol{\alpha}; N, N^\eta)| \ll (HN)^{1+\varepsilon}\left(N^{\Delta_s-(k_1-\phi)/2}\right)^{1/(2s^2)} + HN^{1/2}.$$

But $2\Delta_s - (k_1 - \phi) = \phi - 4s^2\alpha(\mathbf{k})$, whence $\sum_{1\le h\le H} |f_{\mathbf{k}}(h\boldsymbol{\alpha}; N, N^\eta)| = o(N)$. Moreover (see, for example, [19], Lemma 5.3) one has $\mathrm{card}(\mathcal{A}(N, N^\eta)) \gg_\eta N$, and so

$$\sum_{1\le h\le H} |f_{\mathbf{k}}(h\boldsymbol{\alpha}; N, N^\eta)| = o\left(\sum_{n\in\mathcal{A}(N,N^\eta)} 1\right).$$

It therefore follows from [5], Lemma 5, that

$$\min_{1\le n\le N} \|\alpha_1 n^{k_1} + \cdots + \alpha_t n^{k_t}\| < H^{-1} = N^{\phi-\alpha(\mathbf{k})},$$

and the theorem follows in the first case on taking a sequence of q going to infinity.

In the case in which the α_i are in rational ratio, one may of course conclude as above that for infinitely many $n \in \mathbb{N}$ one has $\|\alpha_1 n^{k_1} + \cdots + \alpha_t n^{k_t}\| \le n^{\varepsilon-\alpha(\mathbf{k})}$. However, if the k_i are suitably distributed one may obtain a sharper conclusion. We write $\boldsymbol{\alpha} = \alpha\mathbf{b}$ with $\mathbf{b} \in \mathbb{Z}^t$ and $\alpha \in \mathbb{R}$. Let a/q be a convergent to the continued fraction of α, so that $a \in \mathbb{Z}$, $q \in \mathbb{N}$, $(a,q) = 1$ and $|\alpha - a/q| \le q^{-2}$. We define $\beta(\mathbf{k})$ as in the statement of Theorem 8, with s the value of the parameter corresponding to the maximum. Let ϕ be any real number with $4\varepsilon < \phi < \beta(\mathbf{k})$, and define the real numbers N and H by $N^{k_1-\phi} = q^2$ and $H = N^{\beta(\mathbf{k})-\phi}$. We apply Theorem 8.1 with $\lambda = \frac{1}{2}$ and $m = w = s$ to deduce that

$$\sum_{1\le h\le H} |f_{\mathbf{k}}(h\mathbf{b}\alpha; N, N^\eta)| \ll_{\mathbf{b}} (HN)^{1+\varepsilon}\left(N^{\Delta_s}\Theta\right)^{1/(2s^2)} + HN^{1/2},$$

where

$$\Theta = H^{-1} \min_{\substack{\mathcal{J}\subseteq\{1,\ldots,t\}\\ \mathcal{J}\ne\emptyset}} \prod_{i\in\mathcal{J}} Hq^\varepsilon\left(q^{-1} + N^{-k_i/2} + qH^{-1}N^{-k_i}\right)$$

$$\ll N^{\varepsilon-(k_1-\phi)/2} \min_{2\le r\le t} \prod_{i=2}^r N^{\frac{1}{2}(k_1-\phi)-k_i} \ll N^{t\phi-\mathcal{K}/2}.$$

But $2\Delta_s - (\mathcal{K} - 2t\phi) = 2t\phi - 4s^2\beta(\mathbf{k})$, whence $\sum_{1\le h\le H} |f_{\mathbf{k}}(h\mathbf{b}\alpha; N, N^\eta)| = o_{\mathbf{b}}(N)$. Thus, as in the first case, we deduce that

$$\min_{1\le n\le N} \|\alpha_1 n^{k_1} + \cdots + \alpha_t n^{k_t}\| < H^{-1} = N^{\phi-\beta(\mathbf{k})},$$

and the theorem follows in the second case on taking a sequence of q going to infinity.

The proof of Corollary 8.1 is routine, and so we can afford to be brief.

The proof of Corollary 8.1. In order to prove the first part of the corollary, we put $u_0 = [\frac{1}{2}tk_1(\log t + \log\log t + 4)] + 1$, and note that in the notation of Theorem 2 we have $u_0 < s_0$. Then by Theorem 2 the exponent Δ_{u_0} is admissible, where

$$\Delta_{u_0} = tk_1 \exp(2 - (\log t + \log\log t + 4)) = e^{-2}k_1/\log t.$$

Then when k_1 is large, the conclusion (1.14) of Theorem 8 holds with

$$\alpha(\mathbf{k}) = \frac{k_1 - 2\Delta_{u_0}}{4u_0^2} = \frac{k_1(1 - 2e^{-2}/\log t)}{t^2k_1^2(\log t + \log\log t + 4)^2} - \varepsilon,$$

and a modest computation reveals that when $t > 1$ one has

$$\alpha(\mathbf{k})^{-1} < t^2k_1(\log t + \log\log t + 7)^2.$$

For the second part of the corollary, we may assume that $k_i = o(\sqrt{k_1})$ for $i \geq 2$. Thus, by the argument of the proof of Theorem 3, we may suppose that the exponent Δ_s is admissible, where

$$\Delta_s = \Delta_{s,k_1} + \sum_{i=2}^{t} k_i = \Delta_{s,k_1} + o(k_1) \quad (s \in \mathbb{N}).$$

But the Corollary to [26], Theorem 2.1 implies that when $s \in \mathbb{N}$, the exponent $\Delta_{s,k_1} = \Delta$ is admissible where Δ is the unique positive solution of the equation $\Delta e^{\Delta/k_1} = k_1e^{1-2s/k_1}$. Then provided that s is bounded by a fixed multiple of k_1, one has that $\Delta_s = (1 + o(1))\Delta$ is admissible. Thus we may apply precisely the argument of §4 of [26] to deduce that (1.14) holds with $\alpha(\mathbf{k})^{-1} = (\gamma + o(1))k_1$. This completes the proof of the corollary.

9. Waring's problem for polynomials, I

In this section we establish the first part of Theorem 9 by using an argument which makes use of diminishing ranges. We defer to the next section the homogeneous argument used to prove the second part of Theorem 9. We suppose that $g(x)$ is a polynomial with rational coefficients taking integral values whenever x is an integer. If g has degree k and weight t, then for some $b_i \in \mathbb{Q}$ $(1 \leq i \leq t)$, and $k_1, \dots, k_t$ satisfying (1.2) and $k_1 = k$, one has $g(x) = \sum_{i=1}^t b_ix^{k_i}$. Let d be the least positive integer such that $db_i \in \mathbb{Z}$ $(1 \leq i \leq t)$, and let $\omega = b_1/|b_1|$. Then it follows that n admits a representation in the form (1.15) if and only if ωdn admits a representation in the form (1.15) with $g(x)$ replaced by $\sum_{i=1}^t (\omega db_i)x^{k_i}$. Consequently, we may assume that the polynomial $g(x)$ takes the form $\sum_{i=1}^t a_ix^{k_i}$ with $a_i \in \mathbb{Z}$ $(1 \leq i \leq t)$ and $a_1 > 0$.

Let k_1 be sufficiently large, let n be a sufficiently large positive integer, and write $P = \big[(n/a_1)^{1/k}\big]$. Let ε be a positive number sufficiently small in terms of k, and let η be a positive number sufficiently small in terms of ε and k. Write $R = P^\eta$, and when i is a natural number write $P_i = 2^{-i}P^{(1-1/k)^{i-1}}$. Take s and u to be integer parameters, bounded in terms of k, to be chosen later. We consider the number, $R_{s,u}(n)$, of solutions of the diophantine equation

$$g(x_1) + g(x_2) + \sum_{i=1}^{2s} g(y_i) + \sum_{j=1}^{2u} g(z_j) = n, \tag{9.1}$$

with

$$1 \leq x_1, x_2 \leq P, \quad P_i < y_i, y_{s+i} \leq 2P_i \quad (1 \leq i \leq s), \tag{9.2}$$

and

$$z_j \in \mathcal{A}(P, R) \quad (1 \leq j \leq 2u). \tag{9.3}$$

Write

$$H(\alpha) = \sum_{1 \leq x \leq P} e(\alpha g(x)), \quad h(\alpha) = \sum_{z \in \mathcal{A}(P,R)} e(\alpha g(z)), \tag{9.4}$$

$$f_i(\alpha) = \sum_{P_i < y \le 2P_i} e(\alpha g(y)) \quad (1 \le i \le s), \tag{9.5}$$

and when I is an integer with $1 \le I \le s$, define also

$$\mathcal{F}_{I,s}(\alpha) = \prod_{i=I}^{s} f_i(\alpha). \tag{9.6}$$

Then by orthogonality,

$$R_{s,u}(n) = \int_0^1 \mathcal{F}_{1,s}(\alpha)^2 H(\alpha)^2 h(\alpha)^{2u} e(-\alpha n) d\alpha. \tag{9.7}$$

Before executing the Hardy-Littlewood dissection, it is convenient to provide an estimate for the mean value $U_{I,s}(P)$, which we define by

$$U_{I,s}(P) = \int_0^1 |\mathcal{F}_{I,s}(\alpha)|^2 \, d\alpha. \tag{9.8}$$

Lemma 9.1. *When $k > 2$, and I and s are natural numbers with $I \le s$, one has*

$$U_{I,s}(P) \ll P^{k\left((1-1/k)^{I-1} - (1-1/k)^s\right)}. \tag{9.9}$$

Proof. This is essentially a standard diminishing ranges argument. We note merely that the Mean Value Theorem implies that when x and y are integers with $x > y \ge X$, and X is sufficiently large, then one has $|g(x) - g(y)| > \frac{1}{2} k a_1 X^{k-1}$.

We now apply the Hardy-Littlewood method. Let $\mathfrak{m}$ denote the set of real numbers $\alpha \in [0,1)$ such that whenever $a \in \mathbb{Z}$, $q \in \mathbb{N}$, $(a,q) = 1$ and $|\alpha - a/q| \le q^{-1} P^{\frac{1}{2} - k}$, then one has $q > a_1 P^{1/2} R$. Then recalling the definition of $\mathfrak{n}_\lambda$, we may apply Theorem 7 to deduce that

$$\sup_{\alpha \in \mathfrak{m}} |h(\alpha)| \ll P^{1 - \tau(\mathbf{k}) + \varepsilon}, \tag{9.10}$$

where when k is large one may take $\tau(\mathbf{k})^{-1} = (1 + o(1)) 2tk (\log k + \log t)$. Thus, since k is sufficiently large, we have $\sup_{\alpha \in \mathfrak{m}} |h(\alpha)| \ll P^{1 - \tau_1(k)}$, where

$$\tau_1(k)^{-1} = 6tk \log k. \tag{9.11}$$

We now fix u in terms of s by setting

$$u = \left[\frac{k(1 - 1/k)^s}{2\tau_1(k)} \right] + 1. \tag{9.12}$$

Thus, on using a trivial estimate for $H(\alpha)$, we conclude from Lemma 9.1 that

$$\begin{aligned} \int_{\mathfrak{m}} \left| \mathcal{F}_{1,s}(\alpha)^2 H(\alpha)^2 h(\alpha)^{2u} \right| d\alpha &\le P^2 \left(\sup_{\alpha \in \mathfrak{m}} |h(\alpha)| \right)^{2u} \int_0^1 |\mathcal{F}_{1,s}(\alpha)|^2 \, d\alpha \\ &\ll \mathcal{T}(P) P^{\Delta}, \end{aligned}$$

where we write

$$\mathcal{T}(P) = \mathcal{F}_{1,s}(0)^2 P^{2u+2-k}, \tag{9.13}$$

and $\Delta = k(1 - 1/k)^s - 2u\tau_1(k) < 0$. On recalling (9.11) and (9.12), we may therefore summarise our deliberations thus far with the following lemma.

Lemma 9.2. *Let s be a natural number, and write $u = [3tk^2 \log k(1-1/k)^s]+1$. Then there exists a positive number δ such that*

$$\int_{\mathfrak{m}} \left| \mathcal{F}_{1,s}(\alpha)^2 H(\alpha)^2 h(\alpha)^{2u} \right| d\alpha \ll \mathcal{T}(P)P^{-\delta}.$$

It remains to determine the contribution of the major arcs $\mathfrak{M} = [0,1) \setminus \mathfrak{m}$. We adopt a pruning procedure, taking W to be a pruning parameter with $2 \le W \le P^{1/4}$ to be chosen later. Define

$$\mathfrak{M}(q,a) = \{\alpha \in [0,1) : |q\alpha - a| \le P^{\frac{1}{2}-k}\}. \tag{9.14}$$

Then $\mathfrak{M}$ is the disjoint union of the $\mathfrak{M}(q,a)$ with $0 \le a \le q \le a_1 P^{1/2}R$ and $(a,q)=1$. Define also

$$\mathfrak{N}(q,a) = \{\alpha \in [0,1) : |q\alpha - a| \le WP^{-k}\}, \tag{9.15}$$

and let $\mathfrak{N}$ denote the union of the $\mathfrak{N}(q,a)$ with $0 \le a \le q \le W$ and $(a,q)=1$. Plainly, for each a and q one has $\mathfrak{N}(q,a) \subseteq \mathfrak{M}(q,a)$ and $\mathfrak{N} \subseteq \mathfrak{M}$. We write, further, $\mathfrak{n} = [0,1) \setminus \mathfrak{N}$, so that $\mathfrak{n} = (\mathfrak{M} \setminus \mathfrak{N}) \cup \mathfrak{m}$.

We record in the following lemma standard estimates for the exponential sum $H(\alpha)$ used in our pruning operation. We define $S_g(q,a)$ as in (1.18), and define also

$$I_g(\beta) = \int_0^P e(\beta g(\gamma))d\gamma. \tag{9.16}$$

Further, when $\alpha \in \mathfrak{M}(q,a)$ we write

$$V(\alpha;q,a) = q^{-1}S_g(q,a)I_g(\alpha - a/q). \tag{9.17}$$

Finally, we define the function $V(\alpha)$ to be $V(\alpha;q,a)$ when $\alpha \in \mathfrak{M}(q,a)$, $0 \le a \le q \le a_1P^{1/2}R$ and $(a,q)=1$, and to be zero otherwise.

Lemma 9.3. *When $\alpha \in \mathfrak{M}(q,a)$, one has*

$$H(\alpha) - V(\alpha;q,a) \ll q + |q\alpha - a|P^k. \tag{9.18}$$

Moreover, when $(a,q)=1$,

$$V(\alpha;q,a) \ll q^{\varepsilon}P\left(q + |q\alpha - a|P^k\right)^{-1/k}. \tag{9.19}$$

Proof. The estimate (9.18) follows from Theorem 7.2 of Vaughan [18], while (9.19) follows from Lemma 4.2 of R. Baker [1] combined with Theorem 7.3 of Vaughan [18], on noting that $g(P) \gg P^k$.

Our pruning operation is facilitated by the following lemma.

Lemma 9.4. *Suppose that s is an integer with $s \ge 2k+2$. Then*

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

and there is a positive number δ such that

$$\int_{\mathfrak{M}\setminus\mathfrak{N}} |H(\alpha)|^s d\alpha \ll P^{s-k}W^{-\delta}.$$

Proof. It follows from Lemma 9.3 that when $\alpha \in \mathfrak{M}$ one has $H(\alpha) = V(\alpha) + O(P^{1/2+\varepsilon})$, whence for positive integral s,

$$|H(\alpha)|^s - |V(\alpha)|^s \ll P^{s/2+\varepsilon} + P^{1/2+\varepsilon}|V(\alpha)|^{s-1}. \tag{9.20}$$

Write $\mathcal{M}$ for either $\mathfrak{M}$ or $\mathfrak{M} \setminus \mathfrak{N}$, and define $Y = Y(\mathcal{M})$ to be 1 when $\mathcal{M} = \mathfrak{M}$, and to be W when $\mathcal{M} = \mathfrak{M} \setminus \mathfrak{N}$. Then it also follows from Lemma 9.3 that whenever $s > k$,

$$\int_{\mathcal{M}} |V(\alpha)|^s d\alpha \ll P^{s-k} \sum_{1 \le q \le P} q^{1-s/k+\varepsilon} \min\{1, (q/Y)^{\frac{s}{k}-1}\},$$

and thus for $s > 2k$,

$$\int_{\mathcal{M}} |V(\alpha)|^s d\alpha \ll P^{s-k}Y^{\varepsilon - 1/k}.$$

The lemma now follows immediately from (9.20).

We are now in a position to establish the pruning lemma.

Lemma 9.5. *When $\mathcal{B} \subseteq [0,1)$, write*

$$I(\mathcal{B}) = \int_{\mathcal{B}} |\mathcal{F}_{1,s}(\alpha)H(\alpha)h(\alpha)^u|^2 \, d\alpha.$$

Then provided that $u \geq k$,

$$I([0,1)) \ll \mathcal{T}(P),$$

and there is a positive number δ such that

$$I(\mathfrak{n}) \ll \mathcal{T}(P)W^{-\delta}.$$

Proof. By applying Hölder's inequality we obtain

$$I(\mathcal{M}) \leq \left(\int_{\mathcal{M}} |\mathcal{F}_{1,s}(\alpha)^2 H(\alpha)^{2u+2}| d\alpha \right)^{\frac{1}{u+1}} \left(\int_0^1 \left|\mathcal{F}_{1,s}(\alpha)^2 h(\alpha)^{2u+2}\right| d\alpha \right)^{\frac{u}{u+1}}, \tag{9.21}$$

where $\mathcal{M}$ is either $\mathfrak{M}$ or $\mathfrak{M} \setminus \mathfrak{N}$. Moreover, by considering the underlying diophantine equations,

$$\int_0^1 \left|\mathcal{F}_{1,s}(\alpha)^2 h(\alpha)^{2u+2}\right| d\alpha \leq \int_0^1 \left|\mathcal{F}_{1,s}(\alpha)^2 H(\alpha)^2 h(\alpha)^{2u}\right| d\alpha = I([0,1)). \tag{9.22}$$

But by Lemma 9.2, for some positive number δ one has

$$I([0,1)) = I(\mathfrak{M}) + I(\mathfrak{m}) \ll I(\mathfrak{M}) + \mathcal{T}(P)P^{-\delta}.$$

We note that by making a trivial estimate in combination with Lemma 9.4, whenever $u \geq k$ one has

$$\int_{\mathfrak{M}} |\mathcal{F}_{1,s}(\alpha)^2 H(\alpha)^{2u+2}| d\alpha \leq \mathcal{F}_{1,s}(0)^2 \int_{\mathfrak{M}} |H(\alpha)|^{2u+2} d\alpha \ll \mathcal{T}(P).$$

Then we deduce from (9.21) and (9.22) that whenever $u \geq k$, one has

$$I([0,1)) \ll (\mathcal{T}(P))^{\frac{1}{u+1}} \left(I([0,1))\right)^{\frac{u}{u+1}} + \mathcal{T}(P)P^{-\delta},$$

whence

$$I([0,1)) \ll \mathcal{T}(P). \tag{9.23}$$

This establishes the first part of the lemma.

Next we substitute (9.22) and (9.23) into (9.21). On applying Lemma 9.4, we deduce that for some positive number δ,

$$I(\mathfrak{M} \setminus \mathfrak{N}) \ll \left(\mathcal{T}(P)W^{-(u+1)\delta}\right)^{\frac{1}{u+1}} (\mathcal{T}(P))^{\frac{u}{u+1}}.$$

Then, again recalling Lemma 9.2, we find that

$$I(\mathfrak{n}) = I(\mathfrak{M} \setminus \mathfrak{N}) + I(\mathfrak{m}) \ll \mathcal{T}(P)W^{-\delta}.$$

This completes the proof of the lemma.

In order to complete the proof of the first estimate of Theorem 9, we require estimates for the exponential sum $h(\alpha)$ on suitable pruned major arcs $\mathfrak{N}$. Let $\rho(x)$ denote Dickman's function, defined for real x by

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

We define, further,

$$J_g(\beta) = \int_R^P \rho\left(\frac{\log \gamma}{\log R}\right) e(\beta g(\gamma)) d\gamma, \tag{9.24}$$

and when $\alpha \in \mathfrak{N}(q,a)$ we write

$$W(\alpha; q, a) = q^{-1} S_g(q,a) J_g(\alpha - a/q). \tag{9.25}$$

We require the estimates contained in the following lemmata.

Lemma 9.6. *Suppose that $R = P^\eta$ with η a sufficiently small positive number. Then when $q \le R$, $(a,q)=1$ and $\alpha \in \mathfrak{N}(q,a)$, one has*

$$h(\alpha) - W(\alpha;q,a) \ll \frac{P}{\log P}(q + P^k|q\alpha - a|).$$

Proof. This is immediate from Lemma 8.5 of [22].

Lemma 9.7. *One has*

$$J_g(\beta) \ll P(1+P^k|\beta|)^{-1/k}.$$

Proof. This follows from Lemma 8.6 of [22], on recalling that $g(P) \gg P^k$.

On applying the triangle inequality together with Lemma 9.5, we deduce that for some positive number δ,

$$\begin{aligned}\int_{\mathfrak{M}\setminus\mathfrak{N}} \mathcal{F}_{1,s}(\alpha)^2 H(\alpha)^2 h(\alpha)^{2u} e(-\alpha n)\,d\alpha &\ll \int_{\mathfrak{M}\setminus\mathfrak{N}} \left|\mathcal{F}_{1,s}(\alpha)^2 H(\alpha)^2 h(\alpha)^{2u}\right| d\alpha \\ &\ll \mathcal{T}(P) W^{-\delta}.\end{aligned}$$

Then by (9.7),

$$R_{s,u}(n) = \int_{\mathfrak{N}} \mathcal{F}_{1,s}(\alpha)^2 H(\alpha)^2 h(\alpha)^{2u} e(-\alpha n)\,d\alpha + O\left(\mathcal{T}(P)W^{-\delta}\right).$$

We now obtain an asymptotic formula for the contribution of the pruned major arcs. Write

$$I_{g,i}(\beta) = \int_{P_i}^{2P_i} e(\beta g(\gamma))\,d\gamma \quad (1 \le i \le s).$$

Lemma 9.8. *Suppose that $W \le R$. Then*

$$\int_{\mathfrak{N}} \mathcal{F}_{1,s}(\alpha)^2 H(\alpha)^2 h(\alpha)^{2u} e(-n\alpha)\,d\alpha - \sum_{1\le q\le W} S(q) J^*(q,P,W) \ll \mathcal{T}(P) W^3/\log P,$$

where

$$S(q) = \sum_{\substack{a=1\\(a,q)=1}}^{q} \left(q^{-1} S_g(q,a)\right)^{2s+2u+2} e(-na/q),$$

and

$$J^*(q,P,W) = \mathcal{F}_{2,s}(0)^2 \int_{-q^{-1}WP^{-k}}^{q^{-1}WP^{-k}} I_g(\beta)^2 I_{g,1}(\beta)^2 J_g(\beta)^{2u} e(-n\beta)\,d\beta.$$

Proof. Suppose that $0 \le a \le q \le W$, $(a,q)=1$ and $\alpha \in \mathfrak{N}(q,a)$. Then by Lemma 9.3,

$$H(\alpha) - V(\alpha;q,a) \ll W, \tag{9.26}$$

and by Lemma 9.6, when $W \le R$,

$$h(\alpha) - W(\alpha;q,a) \ll WP/\log P. \tag{9.27}$$

Meanwhile Theorem 7.2 of Vaughan [18] shows that under the same circumstances,

$$f_i(\alpha) - q^{-1} S_g(q,a) I_{g,i}(\alpha - a/q) \ll W \quad (1 \le i \le s). \tag{9.28}$$

We note further that when $2 \le i \le s$, and $|\beta| \le q^{-1}WP^{-k}$, one has

$$I_{g,i}(\beta) = \int_{P_i}^{2P_i} e(\beta g(\gamma))\,d\gamma = P_i + O(q^{-1}WP_i^{k+1}/P^k),$$

whence for some positive number δ,

$$I_{g,i}(\beta) = P_i + O(P^{-\delta}). \tag{9.29}$$

On combining (9.26)-(9.29), we deduce that when $0 \le a \le q \le W$, $(a,q)=1$ and $\alpha \in \mathfrak{N}(q,a)$, one has

$$\mathcal{F}_{1,s}(\alpha)^2 H(\alpha)^2 h(\alpha)^{2u} e(-\alpha n) - \widetilde{S}(q,a)\widetilde{J}(q,P,W;\alpha - a/q) \ll \mathcal{F}_{1,s}(0)^2 P^{2u+2} W/(\log P),$$

where

$$\widetilde{S}(q,a) = \left(q^{-1} S_g(q,a)\right)^{2s+2u+2} e(-na/q),$$

and

$$\widetilde{J}(q,P,W;\beta) = I_g(\beta)^2 I_{g,1}(\beta)^2 J_g(\beta)^{2u} \left(\prod_{i=2}^{s} P_i\right)^2 e(-n\beta).$$

The lemma follows on integrating each side of the last inequality with respect to α over the set $\mathfrak{N}$.

Lemma 9.9. *Let*

$$J(n) = \mathcal{F}_{2,s}(0)^2 \int_{-\infty}^{\infty} I_g(\beta)^2 I_{g,1}(\beta)^2 J_g(\beta)^{2u} e(-n\beta) d\beta.$$

Then when $u > k-2$, the singular integral $J(n)$ converges absolutely, $J(n) \ll \mathcal{T}(P)$, and

$$J^*(q,P,W) - J(n) \ll (q/W)^{1/k} \mathcal{T}(P).$$

Proof. This is a standard consequence of the estimate $I_g(\beta) \ll P(1+P^k|\beta|)^{-1/k}$ which follows from Theorem 7.3 of Vaughan [18], together with Lemma 9.7 (see the proof of Lemma 10.2 of [22] for a similar analysis of a two dimensional case).

Lemma 9.10. *There is a positive number $\mathcal{C}$ such that $J(n) \ge \mathcal{C}\mathcal{T}(P)$.*

Proof. On making a change of variable, we have

$$I_g(\beta) = P \int_0^1 e(\beta g(\gamma P)) d\gamma,$$

$$I_{g,1}(\beta) = P \int_{1/2}^1 e(\beta g(\gamma P)) d\gamma,$$

and

$$J_g(\beta) = P \int_{R/P}^1 \rho\left(\frac{\log P}{\log R} + \frac{\log \gamma}{\log R}\right) e(\beta g(\gamma P)) d\gamma.$$

Thus, on making a second change of variables, we deduce that

$$J(n) = \mathcal{C}(P)\mathcal{T}(P),$$

where

$$\mathcal{C}(P) = \int_{-\infty}^{\infty} \int_{\mathcal{B}} \mathcal{P}(\boldsymbol{\gamma}) e(\beta \Psi(\boldsymbol{\gamma}; P)) d\boldsymbol{\gamma} d\beta,$$

$$\Psi(\boldsymbol{\gamma}; P) = P^{-k}\left(g(\gamma_1 P) + \cdots + g(\gamma_{2u+4} P) - n\right),$$

$$\mathcal{P}(\boldsymbol{\gamma}) = \prod_{i=1}^{2u} \rho(\log(P\gamma_i)/\log R),$$

and

$$\mathcal{B} = (R/P, 1]^{2u} \times [0,1]^2 \times [\tfrac{1}{2}, 1]^2.$$

We note that in view of the absolute convergence of $J(n)$, the contribution to $J(n)$ from the region with $|\beta| > P^{1/3}$ tends to zero as $P \to \infty$.

We observe that $\mathcal{P}(\boldsymbol{\gamma})$ behaves as a positive weight function for each $\boldsymbol{\gamma} \in \mathcal{B}$. Thus a standard application of Fourier's integral formula reveals that

$$\mathcal{C}(P) \geq \int_{-\infty}^{\infty} \int_{\mathcal{B}'} \mathcal{P}(\boldsymbol{\gamma}) e(\beta \Psi(\boldsymbol{\gamma}; P)) d\boldsymbol{\gamma} d\beta,$$

where $\mathcal{B}' = [\frac{1}{2}, 1]^{2u+4}$. Further, since $R = P^\eta$ it follows that when $\frac{1}{2} \leq \gamma \leq 1$,

$$\begin{aligned}\rho(\log(P\gamma)/\log R) &= \rho(\eta^{-1} + \eta^{-1} \log \gamma / \log P) \\ &= \rho(\eta^{-1}) + O(1/\log P),\end{aligned}$$

and hence

$$\mathcal{P}(\boldsymbol{\gamma}) = \rho(\eta^{-1})^{2u} + O(1/\log P).$$

Moreover,

$$e(\beta \Psi(\boldsymbol{\gamma}; P)) = e(\beta a_1(\gamma_1^k + \cdots + \gamma_{2u+4}^k - 1)) + O(\beta P^{-1}).$$

On recalling the concluding remark of the previous paragraph, moreover, we may restrict attention to the situation where $|\beta| \leq P^{1/3}$, in which case the contribution to $\mathcal{C}(P)$ arising from the latter error term is $O(P^{-1/3})$. We therefore deduce that

$$\mathcal{C}(P) \geq \rho(\eta^{-1})^{2u} \mathcal{D} + o(1),$$

where

$$\mathcal{D} = \int_{-\infty}^{\infty} \int_{\mathcal{B}'} e(\beta a_1(\gamma_1^k + \cdots + \gamma_{2u+4}^k - 1)) d\boldsymbol{\gamma} d\beta.$$

A further standard application of Fourier's integral formula reveals that $\mathcal{D}$ is positive. This completes the proof of the lemma.

Lemma 9.11. *The function $S(q)$ is multiplicative. Moroever, when $s + u \geq k$ one has*

$$\sum_{1 \leq q \leq W} q^{1/k} |S(q)| \ll W^\varepsilon.$$

Proof. This is standard (see, for example, Chapters 2 and 4 of Vaughan [18], and the proof of Lemma 10.6 of [22] for a similar argument).

Lemma 9.12. *Suppose that $s + u \geq k$. Then for some $\sigma > 0$ one has*

$$\sum_{1 \leq q \leq W} S(q) J^*(q, P, W) = J(n) \mathfrak{S}(W) + O(W^{-\sigma} \mathcal{T}(P)),$$

where

$$\mathfrak{S}(W) = \sum_{1 \leq q \leq W} S(q).$$

Further, $\mathfrak{S}(n) = \sum_{q=1}^{\infty} S(q)$ converges absolutely, and $\mathfrak{S}(W) - \mathfrak{S}(n) \ll W^{-\tau}$ for some $\tau > 0$.

Proof. This is standard using Lemma 9.11 (see the proof of Lemma 10.7 of [22] for a similar argument).

Combining the conclusions of Lemmata 9.8, 9.9, 9.10 and 9.12, we deduce that when $s+u \geq k$, there are positive numbers $\mathcal{C}$ and σ such that whenever $W \leq R$, one has

$$R_{s,u}(n) \geq \mathcal{C} \mathfrak{S}(n) \mathcal{T}(P) + O((W^3/\log P + W^{-\sigma}) \mathcal{T}(P)).$$

On taking W to be a suitable power of $\log P$, therefore, one finds that

$$R_{s,u}(n) \geq (\mathfrak{S}(n) + o(1)) \mathcal{C} \mathcal{T}(P).$$

Then given a positive number δ, all sufficiently large numbers n with $\mathfrak{S}(n) > \delta$ satisfy the condition

$$R_{s,u}(n) > (\delta + o(1))\mathcal{C}\mathcal{T}(P) \to \infty$$

as $n \to \infty$. Consequently, recalling the definition of $\overline{G}(g)$,

$$\overline{G}(g) \leq 2s + 2u + 2 = 2s + 2[3tk^2 \log k(1 - 1/k)^s] + 4.$$

On taking $s = [k(\log k + \log t + \log\log k)] + 1$, we have

$$tk^2 \log k(1 - 1/k)^s \leq tk^2 \log k \exp(-(\log k + \log t + \log\log k)) \leq k,$$

whence

$$\overline{G}(g) \leq 2k(\log k + \log t + \log\log k + O(1)).$$

This completes the proof of the first part of Theorem 9.

10. Waring's problem for polynomials, II

In this section we establish the second part of Theorem 9 by using a homogeneous argument. We adopt the same notation and conventions as in §9, but now consider the number, $r_{s,u}(n)$, of solutions of the diophantine equation

$$g(x_1) + g(x_2) + \sum_{i=1}^{2s+2u} g(y_i) = n, \tag{10.1}$$

with

$$1 \leq x_1, x_2 \leq P, \quad y_i \in \mathcal{A}(P, R) \quad (1 \leq i \leq 2s + 2u). \tag{10.2}$$

Recalling (9.4) it follows from orthogonality that

$$r_{s,u}(n) = \int_0^1 H(\alpha)^2 h(\alpha)^{2s+2u} e(-\alpha n) d\alpha. \tag{10.3}$$

Before executing the Hardy-Littlewood dissection, we provide an estimate for the mean value $V_w(P, R)$, which we define by

$$V_w(P, R) = \int_0^1 |h(\alpha)|^{2w} d\alpha.$$

Lemma 10.1. *For each positive integer w, one has*

$$V_w(P, R) \ll P^\kappa S_{w,\mathbf{k}}(P, R),$$

where $\kappa = \sum_{i=2}^t k_i$.

Proof. We note that by orthogonality, $V_w(P, R)$ counts the number of solutions of the diophantine system

$$\sum_{i=1}^{w} (g(x_i) - g(y_i)) = 0,$$
$$\sum_{i=1}^{w} (x_i^{k_j} - y_i^{k_j}) = h_j \quad (2 \leq j \leq t),$$

with $x_i, y_i \in \mathcal{A}(P, R)$ $(1 \leq i \leq w)$, as h_j $(2 \leq j \leq t)$ varies over all possible integral $(t-1)$-tuples. Then

$$V_w(P, R) = \sum_{|h_2| \leq wP^{k_2}} \cdots \sum_{|h_t| \leq wP^{k_t}} \int_{\mathbb{T}^t} |G(\boldsymbol{\alpha})|^{2w} e(-h_2\alpha_2 - \cdots - h_t\alpha_t) d\boldsymbol{\alpha},$$

where

$$G(\boldsymbol{\alpha}) = \sum_{x \in \mathcal{A}(P,R)} e(\alpha_1 g(x) + \alpha_2 x^{k_2} + \cdots + \alpha_t x^{k_t}).$$

Thus,

$$\begin{aligned} V_w(P,R) &\ll \left(\prod_{i=2}^{t} 2wP^{k_i}\right) \int_{\mathbb{T}^t} |G(\boldsymbol{\alpha})|^{2w} d\boldsymbol{\alpha} \\ &\ll P^{\kappa} \int_{\mathbb{T}^t} |f_{\mathbf{k}}(\boldsymbol{\alpha};P,R)|^{2w} d\boldsymbol{\alpha}, \end{aligned}$$

and the lemma follows immediately.

We now apply the Hardy-Littlewood method. Let $\mathfrak{m}$ be defined as in §9. Then applying Theorem 7, we deduce that

$$\sup_{\alpha\in\mathfrak{m}} |h(\alpha)| \ll P^{1-\nu(\mathbf{k})+\varepsilon},$$

where when k is large one may take

$$\nu(\mathbf{k})^{-1} = (1+o(1))2k_1\left(3t^2+6t\log\log k_1+\log(k_1\dots k_t)\right).$$

Thus, since k is sufficiently large, we deduce that when $t=o(\sqrt{\log(k_1\dots k_t)})$, then we have $\sup_{\alpha\in\mathfrak{m}}|h(\alpha)| \ll P^{1-\nu_1(\mathbf{k})}$, where

$$\nu_1(\mathbf{k})^{-1} = 4k_1\log(k_1\dots k_t). \tag{10.4}$$

Suppose that $\Delta_s=\Delta_{s,\mathbf{k}}$ is an admissible exponent. We fix u in terms of s by setting

$$u = \left[\frac{\Delta_s}{2\nu_1(\mathbf{k})}\right]+1. \tag{10.5}$$

Thus, on using a trivial estimate for $H(\alpha)$, we conclude from Lemma 10.1 that

$$\begin{aligned} \int_{\mathfrak{m}} \left|H(\alpha)^2h(\alpha)^{2s+2u}\right| d\alpha &\le P^2\left(\sup_{\alpha\in\mathfrak{m}}|h(\alpha)|\right)^{2u}\int_0^1 |h(\alpha)|^{2s}d\alpha \\ &\ll P^{2s+2u+2-k-\delta}, \end{aligned}$$

where $\delta = 2u\nu_1(\mathbf{k})-\Delta_s>0$. On recalling (10.4) and (10.5), we may therefore summarise our deliberations thus far with the following lemma.

Lemma 10.2. *Let s be a natural number, and write $u=[2\Delta_s k_1\log(k_1\dots k_t)]+1$. Then there exists a positive number δ such that*

$$\int_{\mathfrak{m}} |H(\alpha)^2h(\alpha)^{2s+2u}|d\alpha \ll P^{2s+2u+2-k-\delta}.$$

In order to determine the contribution of the major arcs $\mathfrak{M}=[0,1)\setminus\mathfrak{m}$, we merely note that the treatment of §9 may be adapted by replacing the function $\mathcal{F}_{1,s}(\alpha)$ by the function taking the value 1 for $\alpha\in[0,1)$. Thus we find that when $s+u\ge k$, there is a positive number $\mathcal{C}$ such that

$$r_{s,u}(n) \ge (\mathfrak{S}(n)+o(1))\,\mathcal{C}P^{2s+2u+2-k},$$

where

$$\mathfrak{S}(n) = \sum_{q=1}^{\infty}\sum_{\substack{a=1\\(a,q)=1}}^{q}\left(q^{-1}S_g(q,a)\right)^{2s+2u+2}e(-an/q).$$

Then given a positive number δ, all sufficiently large numbers n with $\mathfrak{S}(n)>\delta$ satisfy the condition

$$r_{s,u}(n) > (\delta+o(1))\mathcal{C}P^{2s+2u+2-k} \to \infty$$

as $n\to\infty$. Consequently,

$$\overline{G}(g) \le 2s+2u+2 = 2s+2\left[2\Delta_s k\log(k_1\dots k_t)\right]+4. \tag{10.6}$$

But Theorem 3 implies that when

$$s = \left[\tfrac{1}{2}k\left(\log(k_1 \dots k_t) + 3t^2\right)\right] + tk\left[2\log\log k + \log\log(k_1 \dots k_t) + 1\right],$$

one has that Δ_s is admissible, where

$$\Delta_s \le (\log k)^2 \exp\left(-(2\log\log k + \log\log(k_1 \dots k_t))\right) \ll \left(\log(k_1 \dots k_t)\right)^{-1}.$$

Thus when $t = o\left(\sqrt{\log(k_1 \dots k_t)}\right)$, it follows from (10.6) that

$$\overline{G}(g) \le (1 + o(1))k \log(k_1 \dots k_t).$$

This completes the proof of the second assertion of Theorem 9.

When g is d-lite we modify our above argument by noting that the observation (6.21) implies that whenever the exponent $\Delta_{s,k}$ is admissible, then so is $\Delta_{s,\mathbf{k}}$, where

$$\Delta_{s,\mathbf{k}} = \Delta_{s,k} + \sum_{i=2}^{t} k_i \le \Delta_{s,k} + \tfrac{1}{2}d(d+1).$$

But when $v = \left[\frac{1}{2}k(\log k - 2\log\log k)\right] + 1$, it follows from Theorem 2 that the exponent $\Delta_{v,k} = e^3(\log k)^2$ is admissible (a sharper conclusion would follow from the Corollary to Theorem 2.1 of [26]). Thus when $d \le \log k$, one has that $\Delta_{v,\mathbf{k}}$ is admissible, where $\Delta_{v,\mathbf{k}} = e^4(\log k)^2$. Applying Theorem 1, we deduce that when

$$s = \left[\tfrac{1}{2}k(\log k - 2\log\log k)\right] + tk\left[2\log\log k + \log\log(k_1 \dots k_t)\right] + 1,$$

one has that Δ_s is admissible, where

$$\Delta_s \ll (\log k)^2 \exp\left(-2\log\log k - \log\log(k_1 \dots k_t)\right) \ll \left(\log(k_1 \dots k_t)\right)^{-1}.$$

Thus when $t = o\left(\log k / \log\log k\right)$, it follows from (10.6) that

$$\overline{G}(g) \le (1 + o(1))k \log k.$$

This completes the proof of the final assertion of Theorem 9.

References

1. R. C. Baker, *Diophantine inequalities*, Clarendon Press, Oxford, 1986.
2. C.-J. Chen, *On the representation of a natural number as a sum of terms of the form* $x(x+1)\dots(x+k-1)/k!$, Acta Math. Sinica **9** (1959), 264–270.
3. K. B. Ford, *New estimates for mean values of Weyl sums*, Internat. Math. Res. Notices (1995), 155–171.
4. G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers, 5th ed.*, Oxford University Press, Oxford, 1979.
5. G. Harman, *Trigonometric sums over primes I*, Mathematika **28** (1981), 249–254.
6. G. Harman, *Small fractional parts of additive forms*, Philos. Trans. Roy. Soc. London Ser. A **345** (1993), 327–338.
7. D. R. Heath-Brown, *On the fractional part of* αn^k, Mathematika **35** (1988), 28–37.
8. L.-K. Hua, *On Waring's problem with polynomial summands*, J. Chinese Math. Soc. **1** (1936), 21–61.
9. L.-K. Hua, *On a generalized Waring problem*, Proc. London Math. Soc. (2) **43** (1937), 161–182.
10. L.-K. Hua, *On Waring's problem*, Quart. J. Math., Oxford Ser. **9** (1938), 199–202.
11. L.-K. Hua, *On a generalized Waring problem. II*, J. Chinese Math. Soc. **2** (1940), 175–191.
12. L.-K. Hua, *Additive Theory of Prime Numbers*, Providence, Rhode Island, 1965.
13. A. A. Karatsuba, *Some arithmetical problems with numbers having small prime divisors. Collection of articles in memory of Jurii Vladimirovic Linnik*, Acta Arith. **27** (1975), 489–492. (Russian)
14. Yu. V. Linnik, *On Weyl's sums*, Mat. Sbornik (Rec. Math.) **12** (1943), 23–39.
15. V. I. Nečaev, *The representation of integers by sums of the form* $x(x+1)\dots(x+n-1)/n!$, Dokl. Akad. Nauk SSSR (N. S.) **64** (1949), 159–162.
16. V. I. Nečaev, Waring's problem for polynomials, vol. 38, Trudy Mat. Inst. Steklov., Izdat. Akad. Nauk SSSR, Moscow, 1951, pp. 190–243.
17. V. I. Nečaev, *On the representation of natural numbers as a sum of terms of the form* $x(x+1)\dots(x+n-1)/n!$, Izv. Akad. Nauk SSSR Ser. Mat. **17** (1953), 485–498.
18. R. C. Vaughan, *The Hardy-Littlewood Method*, Cambridge University Press, Cambridge, 1981.
19. R. C. Vaughan, *A new iterative method in Waring's problem*, Acta Math. **162** (1989), 1–71.

20. R. C. Vaughan and T. D. Wooley, *Further improvements in Waring's problem*, Acta Math. **174** (1995), 147–240.
21. I. M. Vinogradov, *The method of trigonometric sums in the theory of numbers*, Trav. Inst. Steklov, 1947.
22. T. D. Wooley, *On simultaneous additive equations, II*, J. Reine Angew. Math. **419** (1991), 141–198.
23. T. D. Wooley, *Large improvements in Waring's problem*, Annals of Math. **135** (1992), 131–164.
24. T. D. Wooley, *On Vinogradov's mean value theorem*, Mathematika **39** (1992), 379–399.
25. T. D. Wooley, *A note on symmetric diagonal equations*, Number theory with an emphasis on the Markoff spectrum (A. D. Pollington and W. Moran, eds.), Marcel Dekker, 1993, pp. 317–321.
26. T. D. Wooley, *The application of a new mean value therorem to the fractional parts of polynomials*, Acta Arith. **65** (1993), 163–179.
27. T. D. Wooley, *Quasi-diagonal behaviour in certain mean value theorems of additive number theory*, J. Amer. Math. Soc. **7** (1994), 221–245.
28. T. D. Wooley, *New estimates for smooth Weyl sums*, J. London Math. Soc. (2) **51** (1995), 1–13.
29. T. D. Wooley, *Note on simultaneous congruences*, J. Number Theory **58** (1996), 288–297.
30. T. D. Wooley, *On simultaneous additive equations, IV* (in preparation).

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1003, U.S.A.
E-mail address: wooley@math.lsa.umich.edu