Annals of Mathematics **175** (2012), 1575–1627
http://dx.doi.org/10.4007/annals.2012.175.3.12

Vinogradov's mean value theorem via efficient congruencing

By Trevor D. Wooley

Abstract

We obtain estimates for Vinogradov's integral that for the first time approach those conjectured to be the best possible. Several applications of these new bounds are provided. In particular, the conjectured asymptotic formula in Waring's problem holds for sums of s kth powers of natural numbers whenever $s \geqslant 2k^2 + 2k - 3$.

1. Introduction

Exponential sums of large degree play a prominent role in the analysis of problems spanning the analytic theory of numbers, and in consequence the estimation of their mean values is of central significance. Some seventy-five years ago, I. M. Vinogradov [32] obtained new estimates for such mean values by exploiting the translation-dilation invariance of associated systems of Diophantine equations. Thereby, he was able to derive powerful new estimates for exponential sums going well beyond those made available via the differencing methods of Weyl and van der Corput. Decisive progress followed in such topics as Waring's problem, the zero-free region for the Riemann zeta function and the distribution modulo 1 of polynomial sequences (see [33], [34] and [35]). Following a decade or so of technical improvement, Vinogradov's mean value theorem evolved into a form little different from that familiar to present day workers, one which for problems of degree d falls short of the strength expected by a factor of order $\log d$. In this paper we obtain significant improvements in estimates associated with Vinogradov's mean value theorem, coming within a stone's throw of the sharpest possible bounds. As we explain in due course, progress of a similar scale may now be realised in numerous allied problems.

In order to describe our conclusions, we must introduce some notation. When k is a natural number and $\boldsymbol{\alpha} \in \mathbb{R}^k$, we consider the exponential sum

$$f_k(\boldsymbol{\alpha}; X) = \sum_{1 \leqslant x \leqslant X} e(\alpha_1 x + \cdots + \alpha_k x^k), \tag{1.1}$$

Supported by a Royal Society Wolfson Research Merit Award.

where $e(z)$ denotes $e^{2\pi iz}$. It follows from orthogonality that, for natural numbers s, the mean value

$$J_{s,k}(X) = \int_{[0,1)^k} |f_k(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \tag{1.2}$$

counts the number of integral solutions of the system of equations

$$x_1^j + \cdots + x_s^j = y_1^j + \cdots + y_s^j \quad (1 \leqslant j \leqslant k), \tag{1.3}$$

with $1 \leqslant x_i, y_i \leqslant X$ $(1 \leqslant i \leqslant s)$. Motivated by a heuristic application of the circle method, it is widely expected that whenever $\varepsilon > 0$, one should have[1]

$$J_{s,k}(X) \ll X^{\varepsilon}(X^s + X^{2s-\frac{1}{2}k(k+1)}). \tag{1.4}$$

Indeed, the discussion surrounding [30, eq. (7.5)] supplies an ε-free version of such a conjecture for $k > 2$. The corresponding lower bound

$$J_{s,k}(X) \gg X^s + X^{2s-\frac{1}{2}k(k+1)}, \tag{1.5}$$

meanwhile, is easily established (see [30, eq. (7.4)]). The main conclusion of this paper, the proof of which we complete in Section 7, is that the estimate (1.4) holds whenever $s \geqslant k(k+1)$.

THEOREM 1.1. *Suppose that s and k are natural numbers with $k \geqslant 2$ and $s \geqslant k(k+1)$. Then for each $\varepsilon > 0$, one has $J_{s,k}(X) \ll X^{2s-\frac{1}{2}k(k+1)+\varepsilon}$.*

If valid, the conjectured bound (1.4) would imply a conclusion of the same shape as that of Theorem 1.1, provided only that $s \geqslant \frac{1}{2}k(k+1)$. In some sense, therefore, Theorem 1.1 comes within a factor 2 of the best possible result of its type. For additive Diophantine systems of large degree k, this is the first occasion on which a conclusion so close to the best possible has been established, for in all previous results one misses the conjectured bounds by a factor of order $\log k$.

A comparison with previous results on Vinogradov's mean value theorem deserves a discussion in two parts. The original method of Vinogradov [32] for estimating $J_{s,k}(X)$ was refined by means of the p-adic argument of Linnik [19] and achieved its most polished form in the work of Karatsuba [18] and Stechkin [26]. Thus, for each natural number s with $s \geqslant k$, one has a bound of the shape

$$J_{s,k}(X) \leqslant D(s,k)X^{2s-\frac{1}{2}k(k+1)+\eta_{s,k}}, \tag{1.6}$$

where $D(s,k)$ is independent of X and $\eta_{s,k} = \frac{1}{2}k^2(1-1/k)^{[s/k]} \leqslant k^2e^{-s/k^2}$. For large integers k, the exponent $\eta_{s,k}$ is appreciably smaller than $1/k$ as soon

[1]Here and throughout, implicit constants in Vinogradov's notation $\ll$ and $\gg$ depend at most on s, k and ε, unless otherwise indicated.

as $s \geqslant 3k^2(\log k + \log\log k)$. When s is sufficiently large in terms of k, this observation permits the proof of an asymptotic formula of the shape

$$J_{s,k}(X) \sim \mathcal{C}(s,k)X^{2s-\frac{1}{2}k(k+1)}, \tag{1.7}$$

wherein $\mathcal{C}(s,k)$ is a positive number depending at most on s and k. Note that the positivity of $\mathcal{C}(s,k)$ is a consequence of the lower bound (1.5). Let $V(k)$ denote the least natural number s for which the anticipated relation (1.7) holds. Then this classical version of Vinogradov's mean value theorem leads to the upper bound $V(k) \leqslant 3k^2(\log k + O(\log\log k))$ (see [30, Th. 7.4]).

The author's thesis work [36], [37] on repeated efficient differencing methods led to sizeable improvements in the conclusions reported in the last paragraph. Roughly speaking, the upper bound (1.6) was established with $\eta_{s,k} \approx k^2e^{-2s/k^2}$ for $s \leqslant k^2\log k$ and with $\eta_{s,k} \approx (\log k)^4e^{-3s/(2k^2)}$ for $s > k^2\log k$ (see [37, Th. 1.2] for a precise statement). In the range critical in applications, the rate of decay of $\eta_{s,k}$ with respect to s stemming from this progress is twice that previously available. As a consequence of these developments, we established that $V(k) \leqslant k^2(\log k + 2\log\log k + O(1))$ (see [42, Th. 3]). We are now able to improve matters significantly.

Define the singular series

$$\mathfrak{S}(s,k) = \sum_{q=1}^{\infty} \sum_{\substack{a_1=1 \\ (a_1,\dots,a_k,q)=1}}^{q} \cdots \sum_{a_k=1}^{q} \Big| q^{-1} \sum_{r=1}^{q} e((a_1r + \cdots + a_kr^k)/q) \Big|^{2s} \tag{1.8}$$

and the singular integral

$$\mathfrak{J}(s,k) = \int_{\mathbb{R}^k} \Big| \int_0^1 e(\beta_1\gamma + \cdots + \beta_k\gamma^k)\,\mathrm{d}\gamma \Big|^{2s} \mathrm{d}\boldsymbol{\beta}. \tag{1.9}$$

It transpires that the positive number $\mathcal{C}(s,k)$ occurring in the putative asymptotic formula (1.7) is then given by $\mathcal{C}(s,k) = \mathfrak{S}(s,k)\mathfrak{J}(s,k)$. In Section 9 we establish the asymptotic formula (1.7) for $s \geqslant k^2 + k + 1$.

THEOREM 1.2. *When $k \geqslant 3$, one has $V(k) \leqslant k^2 + k + 1$.*

The lower bound (1.5) implies that the asymptotic formula (1.7) cannot hold for $s < \frac{1}{2}k(k+1)$. The condition on s imposed in Theorem 1.2 is therefore only a factor 2 away from the best possible conclusion of its type.

The estimate recorded in Theorem 1.1 also leads to improvements in available bounds relating to Tarry's problem. When h, k and s are positive integers with $h \geqslant 2$, consider the Diophantine system

$$\sum_{i=1}^{s} x_{i1}^j = \sum_{i=1}^{s} x_{i2}^j = \ldots = \sum_{i=1}^{s} x_{ih}^j \quad (1 \leqslant j \leqslant k). \tag{1.10}$$

Let $W(k,h)$ denote the least natural number s having the property that the simultaneous equations (1.10) possess an integral solution $\mathbf{x}$ with

$$\sum_{i=1}^{s} x_{iu}^{k+1} \neq \sum_{i=1}^{s} x_{iv}^{k+1} \quad (1 \leqslant u < v \leqslant h).$$

The problem of estimating $W(k,h)$ was investigated extensively by E. M. Wright and L.-K. Hua (see [14], [15], [44]), and very recently upper bounds for $W(k,h)$ have played a role in work of Croot and Hart [8] on the sum-product conjecture. L.-K. Hua was able to show that $W(k,h) \leqslant k^2(\log k + O(1))$ for $h \geqslant 2$, a conclusion improved by the present author when $h = 2$ with the bound $W(k,2) \leqslant \frac{1}{2}k^2(\log k + \log\log k + O(1))$ (see [42, Th. 1]). We improve both estimates in Section 9.

THEOREM 1.3. *When h and k are natural numbers with $h \geqslant 2$ and $k \geqslant 2$, one has $W(k,h) \leqslant k^2 + k - 2$.*

For small values of k, although nothing explicit appears to be available in the literature, some improvement would appear to be possible. It is a simple exercise to show that $W(2,h) = 3$ for $h \geqslant 2$, for example, and the methods of Section 9 of this paper combine with the estimates of [17, Chap. V.5] to confirm that $4 \leqslant W(3,h) \leqslant 8$. On the other hand, explicit numerical examples are available[2] that may be applied to show that $W(k,2) = k+1$ for $2 \leqslant k \leqslant 10$ and $k = 12$.

Next we discuss the asymptotic formula in Waring's problem. When s and k are natural numbers, we denote by $R_{s,k}(n)$ the number of representations of the natural number n as the sum of s kth powers of positive integers. A heuristic application of the circle method suggests that for $k \geqslant 3$ and $s \geqslant k+1$, one should have

$$R_{s,k}(n) = \frac{\Gamma(1+1/k)^s}{\Gamma(s/k)}\mathfrak{S}_{s,k}(n)n^{s/k-1} + o(n^{s/k-1}), \tag{1.11}$$

where

$$\mathfrak{S}_{s,k}(n) = \sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \Big(q^{-1}\sum_{r=1}^{q} e(ar^k/q)\Big)^s e(-na/q). \tag{1.12}$$

Under modest congruence conditions, one has $1 \ll \mathfrak{S}_{s,k}(n) \ll n^{\varepsilon}$, and thus the conjectural relation (1.11) may be interpreted as an honest asymptotic formula (see [30, §§4.3, 4.5 and 4.6] for details). Let $\widetilde{G}(k)$ denote the least integer t with the property that, for all $s \geqslant t$, and all sufficiently large natural numbers

[2]See the website http://euler.free.fr/eslp/eslp.htm.

n, one has the asymptotic formula (1.11). As a consequence of Theorem 1.1, we derive the new upper bound for $\widetilde{G}(k)$ presented in the following theorem.

THEOREM 1.4. *When $k \geqslant 2$, one has $\widetilde{G}(k) \leqslant 2k^2 + 2k - 3$.*

The first to obtain a bound for $\widetilde{G}(k)$ were Hardy and Littlewood [11], who established the bound $\widetilde{G}(k) \leqslant (k-2)2^{k-1} + 5$. The sharpest bounds currently available for smaller values of k are $\widetilde{G}(k) \leqslant 2^k$ $(k = 3, 4, 5)$, due to Vaughan [28], [29], and $\widetilde{G}(k) \leqslant \frac{7}{8}2^k$ $(k \geqslant 6)$, due to Boklan [5]. For larger values of k, the story begins with Vinogradov [32], who showed that $\widetilde{G}(k) \leqslant 183k^9(\log k + 1)^2$. By 1949, Hua [16] had shown that $\widetilde{G}(k) \leqslant (4+o(1))k^2 \log k$. This upper bound was improved first by the author [37] to $\widetilde{G}(k) \leqslant (2 + o(1))k^2 \log k$, and most recently by Ford [9] to $\widetilde{G}(k) \leqslant (1 + o(1))k^2 \log k$. The latter two authors, Parsell [23], and most recently Boklan and Wooley [6], have also computed explicit upper bounds for $\widetilde{G}(k)$ when $k \leqslant 20$. In particular, one has the bounds $\widetilde{G}(7) \leqslant 112$, $\widetilde{G}(8) \leqslant 224$ due to Boklan [5], and $\widetilde{G}(9) \leqslant 365$, $\widetilde{G}(10) \leqslant 497$, $\widetilde{G}(11) \leqslant 627$, $\widetilde{G}(12) \leqslant 771$, $\widetilde{G}(13) \leqslant 934$, $\widetilde{G}(14) \leqslant 1112$, $\widetilde{G}(15) \leqslant 1307$, $\widetilde{G}(16) \leqslant 1517$, $\widetilde{G}(17) \leqslant 1747$, $\widetilde{G}(18) \leqslant 1992$, $\widetilde{G}(19) \leqslant 2255$, $\widetilde{G}(20) \leqslant 2534$ due to Boklan and Wooley [6]. The conclusion of Theorem 1.4 supersedes all of these previous results for $k \geqslant 7$, establishing that $\widetilde{G}(7) \leqslant 109$, $\widetilde{G}(8) \leqslant 141$, $\widetilde{G}(9) \leqslant 177, \dots, \widetilde{G}(20) \leqslant 837$. Furthermore, the strength of Theorem 1.1 opens new possibilities for transforming estimates for $J_{s,k}(X)$ into bounds for auxiliary mean values suitable for investigating Waring's problem. This is a matter that we shall pursue further elsewhere (see [43]).

We turn next to estimates of Weyl type for exponential sums. Here we present conclusions of two types, one applicable to exponential sums $f_k(\boldsymbol{\alpha}; X)$ defined by (1.1) wherein a single coefficient α_j is poorly approximable and a second applicable when $\boldsymbol{\alpha}$ is poorly approximable as a k-tuple.

THEOREM 1.5. *Let k be an integer with $k \geqslant 2$, and let $\boldsymbol{\alpha} \in \mathbb{R}^k$. Suppose that there exists a natural number j with $2 \leqslant j \leqslant k$ such that, for some $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $(a, q) = 1$, one has $|\alpha_j - a/q| \leqslant q^{-2}$ and $q \leqslant X^j$. Then one has*

$$f_k(\boldsymbol{\alpha}; X) \ll X^{1+\varepsilon}(q^{-1} + X^{-1} + qX^{-j})^{\sigma(k)},$$

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

We remark that the factor X^{ε} in the conclusion of Theorem 1.5 may be replaced by $\log(2X)$ if one increases $\sigma(k)^{-1}$ from $2k(k-1)$ to $2k^2 - 2k + 1$.

THEOREM 1.6. *Let k be an integer with $k \geqslant 2$, and let τ and δ be real numbers with $\tau^{-1} > 4k(k-1)$ and $\delta > k\tau$. Suppose that X is sufficiently large in terms of k, δ and τ, and further that $|f_k(\boldsymbol{\alpha}; X)| \geqslant X^{1-\tau}$. Then there exist integers $q, a_1, \dots, a_k$ such that $1 \leqslant q \leqslant X^{\delta}$ and $|q\alpha_j - a_j| \leqslant X^{\delta - j}$ $(1 \leqslant j \leqslant k)$.*

The conclusion of Theorem 1.5 may be compared, for smaller exponents k, with Weyl's inequality (see [30, Lemma 2.4]). The latter provides an estimate of the same shape as that of Theorem 1.5 in the special case $j = k$, with the exponent 2^{k-1} in place of $2k(k-1)$. The conclusion of Theorem 1.5 is therefore superior to Weyl's inequality for $k \geqslant 8$. Subject to the condition $k \geqslant 6$, Heath-Brown [12] has shown that whenever there exist $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $(a,q) = 1$ and $|\alpha - a/q| \leqslant q^{-2}$, then one has

$$\sum_{1 \leqslant x \leqslant X} e(\alpha x^k) \ll X^{1-\frac{8}{3}2^{-k}+\varepsilon}(X^3q^{-1} + 1 + qX^{3-k})^{\frac{4}{3}2^{-k}}. \tag{1.13}$$

With the same conditions on α, Robert and Sargos [24, Th. 4 et Lemme 7] have shown that for $k \geqslant 8$, one has

$$\sum_{1 \leqslant x \leqslant X} e(\alpha x^k) \ll X^{1-3\cdot 2^{-k}+\varepsilon}(X^4q^{-1} + 1 + qX^{4-k})^{\frac{8}{5}2^{-k}}. \tag{1.14}$$

When $k \geqslant 9$, our conclusions in these special situations are superior to those of Heath-Brown, and those of Robert and Sargos, even for the restricted set of α for which either (1.13) or (1.14) prove superior to Weyl's inequality. Finally, the methods of Vinogradov yield results of the type provided by Theorem 1.5 with the exponent $2k(k-1)$ replaced by $(C + o(1))k^2 \log k$ for suitable values of C. For example, Linnik [19] obtained the permissible value $C = 22400$, Hua [16] obtained $C = 4$, and the sharpest bound available hitherto, due to the author [41], is tantamount to $C = 3/2$. We note also that Wooley [37], Ford [9], Parsell [23], and most recently Boklan and Wooley [6], have computed explicit upper bounds for $\sigma(k)$ when $k \leqslant 20$. The conclusion of Theorem 1.5 is superior to these earlier numerical conclusions in all cases and is transparently sharper for larger values of k by a factor asymptotically of order $\log k$. Similar comments apply to the conclusion of Theorem 1.6, a suitable reference to earlier work being [3, Chaps. 4 and 5].

Our final result concerns the distribution modulo 1 of polynomial sequences. Here, we write $\|\theta\|$ for $\min_{y \in \mathbb{Z}} |\theta - y|$.

THEOREM 1.7. *Let k be an integer with $k \geqslant 2$, and define $\tau(k)$ by $\tau(k)^{-1} = 4k(k-1)$. Then whenever $\boldsymbol{\alpha} \in \mathbb{R}^k$ and N is sufficiently large in terms of k and ε, one has*

$$\min_{1 \leqslant n \leqslant N} \|\alpha_1 n + \alpha_2 n^2 + \cdots + \alpha_k n^k\| < N^{\varepsilon - \tau(k)}.$$

For comparison, R. C. Baker [3, Th. 4.5] provides a similar conclusion in which the exponent $4k(k-1)$ is replaced by $(8 + o(1))k^2 \log k$, a conclusion subsequently improved by the author to $(4 + o(1))k^2 \log k$ (see [37, Cor. 1.3]). For smaller values of k, meanwhile, a conclusion similar to that of Theorem 1.7

is delivered by [3, Th. 5.2], but with the exponent 2^{k-1} in place of $4k(k-1)$. The conclusion of Theorem 1.7 is superior to these earlier results for $k \geqslant 10$.

Given the scale of the improvement in estimates made available via Theorem 1.1, it is natural to enquire whether it is now possible to derive visible improvements in the zero-free region for the Riemann zeta function. The estimate supplied by Theorem 1.1 has the shape $J_{s,k}(X) \leqslant D(k,\varepsilon)X^{2s-\frac{1}{2}k(k+1)+\varepsilon}$ for $s \geqslant k(k+1)$, and the nature of the quantity $D(k,\varepsilon)$ plays a critical role in determining the rate of growth of $|\zeta(\sigma+it)|$ with respect to t when σ is close to 1. It seems clear that, while some numerical improvement will be made available via the methods underlying Theorem 1.1, such improvements will not lead to asymptotically significant improvements in the zero-free region. We refer the reader to the work of Ford [10] for a discussion of recent numerical improvements to which our new results may be expected to contribute.

The arguments that underly our proof of Theorem 1.1, which in a nod to the earlier use of *efficient differencing* we refer to loosely as *efficient congruencing* methods, change little when the setting for Vinogradov's mean value theorem is shifted from $\mathbb{Z}$ to the ring of integers of a number field. A significant feature of our estimates in this respect is that when $s \geqslant k(k+1)$, one is at most a factor X^{ε} away from the truth. In common with Birch's application [4] of Hua's lemma in number fields, this aspect of our estimates makes them robust to variation in the degree of the field extension, since the strength of corresponding Weyl-type estimates for exponential sums no longer plays a significant role in applications. Thus, in any number field, one is able to establish the validity of the Hasse Principle, and of Weak Approximation, for diagonal equations of degree d in $2d^2+2d+1$ or more variables, and moreover one is able to obtain the expected density of rational solutions of such equations. Hitherto, such a conclusion was available via the methods of Birch [4] only for diagonal forms of degree d in 2^d+1 or more variables. In a similar manner, the robustness of the efficient congruencing method permits conclusions to be drawn over function fields, such as $\mathbb{F}_q(t)$, matching in strength what is to be found within this paper. We intend to record the consequences of our methods for such problems in forthcoming work.

Finally, the efficient congruencing operation may be applied with success in a number of multidimensional problems related to Vinogradov's mean value theorem. Thus, the work of Arkhipov, Chubarikov and Karatsuba, and Parsell, on exponential sums in many variables (see [2] and [22], for example) may be improved in a manner no less dramatic than can be seen in the context of the version of Vinogradov's mean value theorem described within this paper. This again is a topic to which we intend to return elsewhere.

The methods underlying our proof of Theorem 1.1 are complicated by the need to control nonsingularity constraints modulo various powers of a prime,

and this control must be exercised within an iterative process a step ahead of its application. This and other complicating factors obscure the key ideas of our argument, and so we have taken the liberty of providing, in Section 2 below, a sketch of the fundamental efficient congruencing process. The reader will also find there an outline of the classical approach of Vinogradov, together with the repeated efficient differencing process. Next, in Section 3, we prepare the notation and basic notions required in our subsequent deliberations. The concern of Section 4 is an estimate for a system of basic congruences, and Section 5 describes the conditioning process required to guarantee appropriate nonsingularity conditions in subsequent steps of the efficient congruencing argument. In Section 6 we discuss the efficient congruencing process itself. We combine these tools in Section 7 with an account of the iterative process, ultimately delivering Theorem 1.1. In Section 8 we turn to our first applications, with the proof of Theorems 1.5, 1.6 and 1.7. Next, in Section 9, we consider Tarry's problem, and establish Theorems 1.2 and 1.3. Finally, in Section 10, we consider the asymptotic formula in Waring's problem and establish Theorem 1.4. We finish in Section 11 by describing a heuristic argument that implies essentially the best possible bound of the shape (1.4).

The author is grateful to the two referees of this paper for useful comments.

2. **A sketch of the efficient congruencing process**

Our goal in this section is to offer an indication of the strategy underlying the efficient congruencing process key to our new bounds. At the same time, it is expedient to introduce some notation of use throughout this paper. In what follows, the letter k denotes a fixed integer exceeding 1, the letter s will be a positive integer and ε denotes a sufficiently small positive number. We take X to be a large real number depending at most on k, s and ε, unless otherwise indicated. In an effort to simplify our analysis, we adopt the following convention concerning the number ε. Whenever ε appears in a statement, either implicitly or explicitly, we assert that the statement holds for each $\varepsilon > 0$. Note that the "value" of ε may consequently change from statement to statement. Finally, we make use of vector notation in a slightly unconventional manner. Thus, we may write $a \leqslant \mathbf{z} \leqslant b$ to denote that $a \leqslant z_i \leqslant b$ for $1 \leqslant i \leqslant t$, we may write $\mathbf{z} \equiv \mathbf{w} \pmod p$ to denote that $z_i \equiv w_i \pmod p$ $(1 \leqslant i \leqslant t)$, or on occasion $\mathbf{z} \equiv \xi \pmod p$ to denote that $z_i \equiv \xi \pmod p$ $(1 \leqslant i \leqslant t)$. Confusion should not arise if the reader interprets similar statements in like manner.

We take this opportunity to highlight our use of an important convention throughout Sections 2–7 and Section 11. Since k is considered fixed, we usually find it convenient to drop explicit mention of k from the exponential sum $f_k(\boldsymbol{\alpha}; X)$ and its mean value $J_{s,k}(X)$, defined in (1.1) and (1.2), respectively.

Although potentially a source of confusion, this manoeuvre removes clutter from notation already burdened by other complexities.

We refer to the exponent λ_s as *permissible* when, for each positive number ε, and for any real number X sufficiently large in terms of s, k and ε, one has $J_s(X) \ll X^{\lambda_s+\varepsilon}$. Define λ_s^* to be the infimum of the set of exponents λ_s permissible for s and k, and then put $\eta_s = \lambda_s^* - 2s + \frac{1}{2}k(k+1)$. Thus, whenever X is sufficiently large in terms of s, k and ε, one has

$$J_s(X) \ll X^{\lambda_s^*+\varepsilon}, \tag{2.1}$$

where

$$\lambda_s^* = 2s - \tfrac{1}{2}k(k+1) + \eta_s. \tag{2.2}$$

Note that, in view of the lower bound (1.5) and the trivial estimate $J_s(X) \leqslant X^{2s}$, one has $0 \leqslant \eta_s \leqslant \frac{1}{2}k(k+1)$ for $s \in \mathbb{N}$. Vinogradov's method employs the translation-dilation invariance of the system (1.3) to bound η_{s+k} in terms of η_s by efficiently engineering a strong congruence condition on the variables.

After Linnik [19], the classical approach to Vinogradov's mean value theorem imposes an initial congruence condition on the variables of the system (1.3) by dividing into congruence classes modulo p for a suitably chosen prime p. Let θ be a positive number with $0 < \theta \leqslant 1/k$, and consider a prime number p with $X^\theta < p \leqslant 2X^\theta$. The existence of such a prime is guaranteed by the Prime Number Theorem, or indeed by weaker results such as Bertrand's Postulate. Next, when c and ξ are nonnegative integers, and $\boldsymbol{\alpha} \in [0,1)^k$, define

$$\mathfrak{f}_c(\boldsymbol{\alpha};\xi) = \sum_{\substack{1\leqslant x\leqslant X\\ x\equiv\xi \pmod{p^c}}} e(\psi(x;\boldsymbol{\alpha})), \tag{2.3}$$

where

$$\psi(x;\boldsymbol{\alpha}) = \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_k x^k. \tag{2.4}$$

An application of Hölder's inequality now leads from (1.1) to the bound

$$|f(\boldsymbol{\alpha};X)|^{2s} = \Big|\sum_{\xi=1}^{p^c} \sum_{\substack{1\leqslant x\leqslant X\\ x\equiv\xi \pmod{p^c}}} e(\psi(x;\boldsymbol{\alpha}))\Big|^{2s} \leqslant (p^c)^{2s-1} \sum_{\xi=1}^{p^c} |\mathfrak{f}_c(\boldsymbol{\alpha};\xi)|^{2s}. \tag{2.5}$$

Let us focus now on the mean value $J_{s+k}(X)$ defined via (1.2). In order to save clutter, when $G : [0,1)^k \to \mathbb{C}$ is measurable, we write

$$\oint G(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha} = \int_{[0,1)^k} G(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha}.$$

On substituting (2.5) into the analogue of (1.2) with s replaced by $s+k$, we find that

$$J_{s+k}(X) \ll X^{2s\theta} \max_{1\leqslant\xi\leqslant p} \oint |f(\boldsymbol{\alpha};X)^{2k}\mathfrak{f}_1(\boldsymbol{\alpha};\xi)^{2s}|\,\mathrm{d}\boldsymbol{\alpha}. \tag{2.6}$$

The mean value on the right-hand side of (2.6) counts the number of integral solutions of the system

$$\sum_{i=1}^{k}(x_i^j-y_i^j)=\sum_{l=1}^{s}((pu_l+\xi)^j-(pv_l+\xi)^j)\quad(1\leqslant j\leqslant k),\tag{2.7}$$

with $1\leqslant \mathbf{x},\mathbf{y}\leqslant X$ and $(1-\xi)/p\leqslant \mathbf{u},\mathbf{v}\leqslant (X-\xi)/p$. But, as a consequence of the Binomial Theorem, the validity of the equations (2.7) implies that

$$\sum_{i=1}^{k}((x_i-\xi)^j-(y_i-\xi)^j)=p^j\sum_{l=1}^{s}(u_l^j-v_l^j)\quad(1\leqslant j\leqslant k),\tag{2.8}$$

whence

$$\sum_{i=1}^{k}(x_i-\xi)^j\equiv\sum_{i=1}^{k}(y_i-\xi)^j\pmod{p^j}\quad(1\leqslant j\leqslant k).\tag{2.9}$$

The congruences (2.9) provide the *efficient* congruence condition mentioned earlier, with the artificial condition modulo p imposed via (2.5) converted into a system of congruence conditions modulo p^j for $1\leqslant j\leqslant k$, as opposed merely to a system of congruence conditions modulo p. Suppose that $\mathbf{x}$ is *well-conditioned*, by which we mean that $x_1,\dots,x_k$ lie in distinct congruence classes modulo p. Then, given an integral k-tuple $\mathbf{n}$, the solutions of the system

$$\sum_{i=1}^{k}(x_i-\xi)^j\equiv n_j\pmod{p}\quad(1\leqslant j\leqslant k),$$

with $1\leqslant \mathbf{x}\leqslant p$, may be lifted uniquely to solutions of the system

$$\sum_{i=1}^{k}(x_i-\xi)^j\equiv n_j\pmod{p^k}\quad(1\leqslant j\leqslant k),$$

with $1\leqslant \mathbf{x}\leqslant p^k$. In this way, the congruences (2.9) essentially imply that

$$\mathbf{x}\equiv\mathbf{y}\pmod{p^k},\tag{2.10}$$

provided that we inflate our estimates by the combinatorial factor $k!$ to account for the multiplicity of solutions modulo p, together with a factor $p^{\frac{1}{2}k(k-1)}$ to account for solutions introduced as one considers for a fixed $\mathbf{n}'$ the possible choices for $\mathbf{n}\pmod{p^k}$ with $n_j\equiv n_j'\pmod{p^j}$ $(1\leqslant j\leqslant k)$.

In the classical argument, one chooses $\theta=1/k$, so that $p^k>X$. Since $1\leqslant \mathbf{x},\mathbf{y}\leqslant X$, one is then forced to conclude from the congruential condition (2.10) that $\mathbf{x}=\mathbf{y}$, and in (2.8) this in turn implies that

$$\sum_{l=1}^{s}(u_l^j-v_l^j)=0\quad(1\leqslant j\leqslant k).$$

The number of solutions of this system with $(1-\xi)/p \leqslant \mathbf{u}, \mathbf{v} \leqslant (X-\xi)/p$ is readily seen to be $O(J_s(X/p))$, and thus the corresponding number of solutions of (2.8) with $\mathbf{x} = \mathbf{y}$ is $O(X^k J_s(X/p))$. Thus, in view of (2.1), one obtains

$$J_{s+k}(X) \ll (X^\theta)^{2s+\frac{1}{2}k(k-1)} X^k J_s(X/p) \ll (X^\theta)^{2s+\frac{1}{2}k(k-1)} X^k (X^{1-\theta})^{\lambda_s^*+\varepsilon}.$$

Since $\theta = 1/k$, it follows from (2.2) that $\lambda_{s+k}^* \leqslant 2(s+k) - \frac{1}{2}k(k+1) + \eta_{s+k}$, with $\eta_{s+k} \leqslant \eta_s(1-1/k)$. On recalling the estimate $J_k(X) \leqslant k!X^k$, stemming from Newton's formulae, the classical bound (1.6) with $\eta_s = \frac{1}{2}k^2(1-1/k)^{[s/k]}$ follows by induction.

Suppose now that we take $\theta < 1/k$, and interpret the condition (2.10) by defining the k-tuple $\mathbf{h}$ by means of the relation $\mathbf{h} = (\mathbf{x}-\mathbf{y})p^{-k}$, so that $\mathbf{x} = \mathbf{y} + \mathbf{h}p^k$. On substituting into (2.8), we obtain the new system

$$\sum_{i=1}^{k} \Psi_j(y_i, h_i, p) = \sum_{l=1}^{s} (u_l^j - v_l^j) \quad (1 \leqslant j \leqslant k), \tag{2.11}$$

where

$$\Psi_j(y,h,p) = p^{-j}((y+hp^k-\xi)^j - (y-\xi)^j) \quad (1 \leqslant j \leqslant k).$$

The number of solutions of the system (2.11) subject to the associated conditions $1 \leqslant \mathbf{y} \leqslant X$, $|\mathbf{h}| \leqslant Xp^{-k}$ and $(1-\xi)/p \leqslant \mathbf{u}, \mathbf{v} \leqslant (X-\xi)/p$, may be reinterpreted by means of an associated mean value of exponential sums. An application of Schwarz's inequality[3] bounds this mean value in terms of $J_s(X/p)$ and a new mean value that counts integral solutions of the system

$$\sum_{i=1}^{k} (\Psi_j(x_i,h,p) - \Psi_j(y_i,h,p)) = \sum_{l=1}^{s} (u_l^j - v_l^j) \quad (1 \leqslant j \leqslant k), \tag{2.12}$$

with variables satisfying similar conditions to those above. We now have the option of repeating the process of imposing an efficient congruence condition on $\mathbf{x}$ and $\mathbf{y}$, much as before, by pushing the variables $\mathbf{u}$ and $\mathbf{v}$ into congruence classes modulo a suitable new prime number ϖ. In this way, one may estimate $J_{s+k}(X)$ by iteratively bounding the number of solutions of a system of type (2.12) by a similar one, wherein the polynomial $\Psi_j(z) = \Psi_j(z,h,p)$ is replaced for $1 \leqslant j \leqslant k$ by one of the shape $\Phi_j(z,g,\varpi) = \varpi^{-j}(\Psi_j(z+g\varpi^k) - \Psi_j(z))$. This *repeated efficient differencing process*, so-called owing to its resemblance to classical Weyl differencing, delivers the more efficient choice of parameter $\theta \approx k/(k^2+\eta_s)$. In the most important range for s, one obtains an estimate roughly of the shape $\eta_{s+k} \leqslant \eta_s(1-2k/(k^2+\eta_s))$, and this yields $\eta_s \approx k^2 e^{-2s/k^2}$ (see [37] for details).

[3]At the prompting of one of the referees, we point out that less pedantic readers may prefer to refer to this inequality as the Cauchy-Schwarz inequality, or arguably more precisely as Bunyakovsky's inequality.

The strategy underlying Vinogradov's method, as seen in both its classical and repeated efficient differencing formulations, is that of transforming an initial congruence condition into a differencing step, with the ultimate aim in (2.7), and its variants such as (2.12), of forcing $2k$ variables to obey a diagonal condition. In this paper we instead view Vinogradov's method as an efficient generator of congruence conditions. Thus, the initial condition modulo p amongst $2s$ variables underlying the mean value $J_{s+k}(X)$ efficiently generates the stronger condition modulo p^k visible in (2.10). Our strategy now is to exploit this condition so as to push $2s$ variables into the same congruence class modulo p^k within a new mean value and efficiently extract from this a fresh congruence condition modulo p^{k^2}. By repeating this process, one extracts successively stronger congruence conditions, and these may be expected to yield successively stronger mean value estimates.

There is a critical detail concerning which we have, thus far, remained silent. We supposed in advance of (2.10) that the k-tuple $\mathbf{x}$ was well-conditioned, and indeed similar assumptions must be made at each point of the repeated efficient differencing process. There are several possible approaches to the challenge of ensuring this well-conditioning of variables, the most straightforward being to preselect the prime so that the bulk of solutions are well-conditioned (see [38] for a transparent application of this idea). The problem of ensuring well-conditioning causes considerable difficulty in the analysis of the efficient congruencing argument in this paper, for our prime is fixed, once and for all, at the outset of our argument. For now we ignore this complication so as to better expose the underlying ideas.

We now outline the repeated efficient congruencing argument. In the first instance, we take $0 < \theta \leqslant 1/k^2$. Observe that, in view of the condition (2.10), one may derive from (2.7) the upper bound

$$J_{s+k}(X) \ll (X^\theta)^{2s+\frac{1}{2}k(k-1)} \max_{1\leqslant \xi \leqslant p} \oint \Big(\sum_{\eta=1}^{p^k} |\mathfrak{f}_k(\boldsymbol{\alpha};\eta)|^2 \Big)^k |\mathfrak{f}_1(\boldsymbol{\alpha};\xi)|^{2s}\, \mathrm{d}\boldsymbol{\alpha}.$$

By Hölder's inequality, therefore, one sees that

$$J_{s+k}(X) \ll (X^\theta)^{2s+\frac{1}{2}k(k-1)} (X^{k\theta})^k \max_{1\leqslant \xi \leqslant p} \max_{1\leqslant \eta \leqslant p^k} I(\xi,\eta), \tag{2.13}$$

where

$$I(\xi,\eta) = \oint |\mathfrak{f}_k(\boldsymbol{\alpha};\eta)^{2k} \mathfrak{f}_1(\boldsymbol{\alpha};\xi)^{2s}|\, \mathrm{d}\boldsymbol{\alpha}.$$

A further application of Hölder's inequality shows that

$$I(\xi,\eta) \leqslant \Big(\oint |\mathfrak{f}_1(\boldsymbol{\alpha};\xi)|^{2s+2k}\, \mathrm{d}\boldsymbol{\alpha} \Big)^{1-k/s} \Big(\oint |\mathfrak{f}_1(\boldsymbol{\alpha};\xi)^{2k} \mathfrak{f}_k(\boldsymbol{\alpha};\eta)^{2s}|\, \mathrm{d}\boldsymbol{\alpha} \Big)^{k/s}. \tag{2.14}$$

Notice that in the second mean value on the right-hand side of (2.14), there is a reversal of rôles of the generating functions $\mathfrak{f}_1(\boldsymbol{\alpha};\xi)$ and $\mathfrak{f}_k(\boldsymbol{\alpha};\eta)$ as compared

to the corresponding mean value defining $I(\xi,\eta)$. As we explain below, it is this manoeuvre that permits repeated application of the congruencing step.

The first integral on the right-hand side of (2.14) counts the number of integral solutions of the system

$$\sum_{i=1}^{s+k}((pu_i+\xi)^j-(pv_i+\xi)^j)=0 \quad (1\leqslant j\leqslant k),$$

with $(1-\xi)/p\leqslant \mathbf{u},\mathbf{v}\leqslant (X-\xi)/p$. An application of the Binomial Theorem shows this to be $O(J_{s+k}(X/p))$. By orthogonality, meanwhile, the second integral is bounded above by the number of solutions of the system

$$\sum_{i=1}^{k}(x_i^j-y_i^j)=\sum_{l=1}^{s}((p^ku_l+\eta)^j-(p^kv_l+\eta)^j) \quad (1\leqslant j\leqslant k), \tag{2.15}$$

with $1\leqslant \mathbf{x},\mathbf{y}\leqslant X$, $\mathbf{x}\equiv\mathbf{y}\equiv\xi \pmod p$ and $(1-\eta)/p^k\leqslant \mathbf{u},\mathbf{v}\leqslant (X-\eta)/p^k$. As in the classical treatment sketched above, it follows as a consequence of the Binomial Theorem that the validity of the equations (2.15) implies that

$$\sum_{i=1}^{k}((x_i-\eta)^j-(y_i-\eta)^j)=p^{jk}\sum_{l=1}^{s}(u_l^j-v_l^j) \quad (1\leqslant j\leqslant k),$$

whence

$$\sum_{i=1}^{k}(x_i-\eta)^j\equiv\sum_{i=1}^{k}(y_i-\eta)^j \pmod{p^{jk}} \quad (1\leqslant j\leqslant k). \tag{2.16}$$

The system (2.16) provides an even more efficient congruence condition than that offered by (2.9), tempered with a slightly diminished return stemming from the fact that the x_i and y_i all lie in the common congruence class ξ modulo p. On the face of it, the latter unequivocally prevents these variables being well-conditioned. However, let us assume for now that $x_1,\dots,x_k$ are distinct modulo p^2, and likewise $y_1,\dots,y_k$. It transpires that on this occasion, one may lift solutions modulo p^2 to solutions modulo p^{k^2}. Indeed, the congruences (2.16) essentially imply that

$$\mathbf{x}\equiv\mathbf{y} \pmod{p^{k^2}}, \tag{2.17}$$

provided that one inserts a compensating factor $k!(p^{k+1})^{\frac{1}{2}k(k-1)}$ into the concomitant estimates. At this point one could repeat the whole process, employing (2.17) to engineer a fresh congruence condition modulo p^{k^3}, then modulo p^{k^4}, and so on. However, in order to illuminate this efficient congruencing argument, we examine instead the consequences of the assumption that $\theta=1/k^2$. In such circumstances, one has $p^{k^2}>X$, and so it follows from (2.17) that

$\mathbf{x} = \mathbf{y}$. Since $\mathbf{x} \equiv \mathbf{y} \equiv \xi \pmod{p}$, the number of possible choices for $\mathbf{x}$ and $\mathbf{y}$ is $O((X/p)^k)$. Substituting into (2.15), we deduce that

$$\oint |\mathfrak{f}_1(\boldsymbol{\alpha};\xi)^{2k}\mathfrak{f}_k(\boldsymbol{\alpha};\eta)^{2s}|\,\mathrm{d}\boldsymbol{\alpha} \ll (X^\theta)^{\frac{1}{2}k(k^2-1)}(X^{1-\theta})^k \oint |\mathfrak{f}_k(\boldsymbol{\alpha};\eta)|^{2s}\,\mathrm{d}\boldsymbol{\alpha} \ll (X^\theta)^{\frac{1}{2}k(k^2-1)}(X^{1-\theta})^k J_s(X/p^k). \tag{2.18}$$

If we now substitute (2.18) into (2.14), we obtain

$$I(\xi,\eta) \ll (J_{s+k}(X/p))^{1-k/s}\Big((X^\theta)^{\frac{1}{2}k(k^2-1)}(X^{1-\theta})^k J_s(X/p^k)\Big)^{k/s},$$

whence, in view of (2.1), it follows from (2.13) that

$$J_{s+k}(X) \ll \Big((X^\theta)^{2s+2k-\frac{1}{2}k(k+1)}(X^{1-\theta})^{\lambda^*_{s+k}+\varepsilon}\Big)^{1-k/s} \times \Big((X^\theta)^{2ks-\frac{1}{2}k(k+1)+2k+\frac{1}{2}k(k^2-1)}(X^{1-\theta})^k(X^{1-k\theta})^{\lambda^*_s+\varepsilon}\Big)^{k/s}.$$

Consequently, from (2.2) we discern the upper bound

$$J_{s+k}(X) \ll (X^{\eta_{s+k}(1-\theta)})^{1-k/s}\Big(X^{-k+k^3\theta}(X^{\eta_s(1-k\theta)})\Big)^{k/s} X^{2s+2k-\frac{1}{2}k(k+1)+\varepsilon}.$$

Recall that $\theta = 1/k^2$. Since $\lambda^*_{s+k} = 2s + 2k - \frac{1}{2}k(k+1) + \eta_{s+k}$ is an infimal exponent, it follows that for a sequence of values of X tending to ∞, one has

$$X^{\eta_{s+k}-\varepsilon} \ll X^\varepsilon (X^{1-1/k^2})^{(1-k/s)\eta_{s+k}}(X^{1-1/k})^{(k/s)\eta_s},$$

whence for each positive number ε, one has

$$\eta_{s+k} \leqslant (1-k/s)(1-1/k^2)\eta_{s+k} + (k/s)(1-1/k)\eta_s + \varepsilon.$$

Noting again the infimal definition of λ^*_{s+k}, we therefore deduce that

$$\eta_{s+k} \leqslant \frac{(1-1/k)\eta_s}{1+(s/k-1)(1/k^2)}. \tag{2.19}$$

Provided that s is no larger than about $k^{5/2}$, a modest computation leads from the iterative relation (2.19) to the upper bound

$$\eta_{s+k} \leqslant (1-s/k^3)\eta_s \leqslant e^{-s/k^3}\eta_s.$$

One therefore sees that η_s is no larger than about $k^2e^{-\frac{1}{2}(s/k^2)^2}$. By comparison with the classical bound $\eta_s \leqslant k^2e^{-s/k^2}$ mentioned following (1.6), one has considerable additional decay in the upper bound for η_s as soon as s is a little larger than k^2. Indeed, even an estimate of this quality would establish, for example, that $\widetilde{G}(k) \ll k^2(\log k)^{1/2}$, greatly improving the bound $\widetilde{G}(k) \leqslant (1+o(1))k^2\log k$ due to Ford [9].

For each natural number N, the pursuit of an N-fold repeated efficient congruencing process delivers bounds with the approximate shape

$$\eta_{s+k} \leqslant \frac{\eta_s}{1+(s/k)^N(1/k^{N+1})}.$$

When $s > k^2$, it is apparent that the upper bound on the right-hand side here converges to zero as N goes to infinity. Such a bound comes close to delivering Theorem 1.1. Two serious obstructions remain. The first is the removal of the assumption throughout that variables are suitably well-conditioned whenever this is essential. Since our auxiliary prime number p is fixed, once and for all, at the opening of our argument, we are forced to engineer well-conditioning directly using this single prime p. Such has the potential to weaken substantially our conclusions, and we are forced to consider a complex iterative process rather difficult to control. The second obstruction is less severe. The condition $s > k^2$ must be replaced by $s = k^2$, and owing to the possibility of ill-conditioned solutions, a direct approach would be successful, at best, only when $s \geqslant k^2 + k$. Once again, therefore, we are forced to negotiate delicate issues associated with a complex iterative process.

3. **Preliminary manoeuvres**

We begin in this section with some notation and definitions of use in our subsequent discussion. Let k be a fixed integer with $k \geqslant 2$, and let δ be a small positive number. We consider a natural number u with $u \geqslant k$, and we put $s = uk$. Our goal is to show that $\lambda^*_{s+k} = 2(s+k) - \frac{1}{2}k(k+1)$, whence $\eta_{s+k} = 0$. In view of the infimal definition of λ^*_{s+k}, there exists a sequence of natural numbers $(X_n)_{n=1}^{\infty}$, tending to infinity, with the property that

$$J_{s+k}(X_n) > X_n^{\lambda^*_{s+k}-\delta} \quad (n \in \mathbb{N}). \tag{3.1}$$

Provided that X_n is sufficiently large, we have also for $X_n^{\delta^2} < Y \leqslant X_n$ the corresponding upper bounds

$$J_t(Y) < Y^{\lambda^*_t+\delta} \quad (t = s, s+k). \tag{3.2}$$

Notice that when $s > k^2$, the trivial inequality $|f(\boldsymbol{\alpha}; X)| \leqslant X$ leads from (1.2) to the upper bound

$$J_{s+k}(X) \leqslant X^{2(s-k^2)} \oint |f(\boldsymbol{\alpha}; X)|^{2k(k+1)}\, \mathrm{d}\boldsymbol{\alpha} \leqslant X^{2(s-k^2)} J_{k(k+1)}(X).$$

It then follows from the above discussion that whenever $s > k^2$, one has $\eta_{s+k} \leqslant \eta_{k(k+1)}$. With an eye toward future applications, we shall continue to consider general values of s with $s \geqslant k^2$ until the very climax of the proof of Theorem 1.1, and only at that point specialise to the situation with $s = k^2$. As we have just shown, the desired conclusion when $s > k^2$ is an easy consequence of this special case. Finally, we take N to be a natural number sufficiently large in terms of s and k, and we put $\theta = \frac{1}{2}(k/s)^{N+1}$. Note that we are at liberty to take δ to be a positive number with $\delta < (Ns)^{-3N}$, so that δ is in particular small compared to θ. We focus now on a fixed element $X = X_n$ of the sequence

(X_n), which we may assume to be sufficiently large in terms of s, k, N and δ, and put $M = X^\theta$. Thus we have $X^\delta < M^{1/N}$.

Let p be a fixed prime number with $M < p \leqslant 2M$ to be chosen in due course. That such a prime exists is a consequence of the Prime Number Theorem. We will find it necessary to consider *well-conditioned* k-tuples of integers belonging to distinct congruence classes modulo a suitable power of p. Denote by $\Xi_c(\xi)$ the set of k-tuples $(\xi_1, \dots, \xi_k)$, with

$$1 \leqslant \xi_i \leqslant p^{c+1} \quad \text{and} \quad \xi_i \equiv \xi \pmod{p^c}\ (1 \leqslant i \leqslant k),$$

and satisfying the property that $\xi_i \equiv \xi_j \pmod{p^{c+1}}$ for no i and j with $1 \leqslant i < j \leqslant k$. In addition, write $\Sigma_k = \{1, -1\}^k$, and consider an element $\boldsymbol{\sigma}$ of Σ_k. Recalling the definition (2.3), we then put

$$\mathfrak{F}_c^{\boldsymbol{\sigma}}(\boldsymbol{\alpha}; \xi) = \sum_{\boldsymbol{\xi} \in \Xi_c(\xi)} \prod_{i=1}^{k} \mathfrak{f}_{c+1}(\sigma_i \boldsymbol{\alpha}; \xi_i). \tag{3.3}$$

Two mixed mean values play leading roles in our arguments. When a and b are nonnegative integers, and $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \Sigma_k$, we define

$$I_{a,b}^{\boldsymbol{\sigma}}(X; \xi, \eta) = \oint |\mathfrak{F}_a^{\boldsymbol{\sigma}}(\boldsymbol{\alpha}; \xi)^2 \mathfrak{f}_b(\boldsymbol{\alpha}; \eta)^{2s}| \, \mathrm{d}\boldsymbol{\alpha} \tag{3.4}$$

and

$$K_{a,b}^{\boldsymbol{\sigma},\boldsymbol{\tau}}(X; \xi, \eta) = \oint |\mathfrak{F}_a^{\boldsymbol{\sigma}}(\boldsymbol{\alpha}; \xi)^2 \mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha}; \eta)^{2u}| \, \mathrm{d}\boldsymbol{\alpha}. \tag{3.5}$$

It is convenient then to put

$$I_{a,b}(X) = \max_{1 \leqslant \xi \leqslant p^a} \max_{1 \leqslant \eta \leqslant p^b} \max_{\boldsymbol{\sigma} \in \Sigma_k} I_{a,b}^{\boldsymbol{\sigma}}(X; \xi, \eta) \tag{3.6}$$

and

$$K_{a,b}(X) = \max_{1 \leqslant \xi \leqslant p^a} \max_{1 \leqslant \eta \leqslant p^b} \max_{\boldsymbol{\sigma}, \boldsymbol{\tau} \in \Sigma_k} K_{a,b}^{\boldsymbol{\sigma},\boldsymbol{\tau}}(X; \xi, \eta). \tag{3.7}$$

Notice here that these mean values depend on our choice of p. However, since we will shortly fix this choice of p once and for all, we suppress mention of this prime when referring to $I_{a,b}(X)$ and $K_{a,b}(X)$.

Our arguments are simplified considerably by making transparent the relationship between various mean values on the one hand and the anticipated magnitude of these mean values on the other. Of course, such a concept may not be well defined, and so we indicate in what follows quite concretely what is intended. We define the *normalised magnitude* of a mean value $\mathcal{M}$ relative to its anticipated size $\mathcal{M}^*$ to be $\mathcal{M}/\mathcal{M}^*$, a quantity we denote by $[[\mathcal{M}]]$. In particular, we define

$$[[J_t(X)]] = \frac{J_{t,k}(X)}{X^{2t - \frac{1}{2}k(k+1)}} \quad (t = s, s+k), \tag{3.8}$$

and when $0 \leqslant a < b$, we define

$$[[I_{a,b}(X)]] = \frac{I_{a,b}(X)}{(X/M^b)^{2s}(X/M^a)^{2k-\frac{1}{2}k(k+1)}},$$

$$[[K_{a,b}(X)]] = \frac{K_{a,b}(X)}{(X/M^b)^{2s}(X/M^a)^{2k-\frac{1}{2}k(k+1)}}. \tag{3.9}$$

Note that the lower bound (3.1) implies that

$$[[J_{s+k}(X)]] > X^{\eta_{s+k}-\delta} \tag{3.10}$$

while the upper bound (3.2) ensures that, whenever $X^{\delta^2} < Y \leqslant X$, one has

$$[[J_t(Y)]] < Y^{\eta_t+\delta} \quad (t = s, s+k). \tag{3.11}$$

Mean values of the exponential sum $\mathfrak{f}_c(\boldsymbol{\alpha};\xi)$ are easily bounded by exploiting the translation-dilation invariance of the solution sets of the system of equations (1.3). The argument is relatively familiar, though we provide details for the sake of completeness.

LEMMA 3.1. *Suppose that c is a nonnegative integer with $c\theta \leqslant 1$. Then for each natural number t, one has*

$$\max_{1\leqslant\xi\leqslant p^c} \oint |\mathfrak{f}_c(\boldsymbol{\alpha};\xi)|^{2t}\, d\boldsymbol{\alpha} \ll_t J_t(X/M^c). \tag{3.12}$$

Proof. Let ξ be an integer with $1 \leqslant \xi \leqslant p^c$. From the definition (2.3) of the exponential sum $\mathfrak{f}_c(\boldsymbol{\alpha};\xi)$, one has

$$\mathfrak{f}_c(\boldsymbol{\alpha};\xi) = \sum_{(1-\xi)/p^c\leqslant y\leqslant(X-\xi)/p^c} e(\psi(p^c y+\xi;\boldsymbol{\alpha})),$$

in which $\psi(z;\boldsymbol{\alpha})$ is given by (2.4). By orthogonality, therefore, one finds that the integral on the left-hand side of (3.12) counts the number of integral solutions of the system of equations

$$\sum_{i=1}^{t}(p^c y_i+\xi)^j = \sum_{i=1}^{t}(p^c z_i+\xi)^j \quad (1 \leqslant j \leqslant k), \tag{3.13}$$

with $0 \leqslant \mathbf{y}, \mathbf{z} \leqslant (X-\xi)/p^c$. An application of the Binomial Theorem shows that the pair $\mathbf{y}, \mathbf{z}$ satisfies (3.13) if and only if it satisfies the system

$$\sum_{i=1}^{t} y_i^j = \sum_{i=1}^{t} z_i^j \quad (1 \leqslant j \leqslant k).$$

Thus, on considering the underlying Diophantine system and recalling (1.1) and (1.2), we find that

$$\begin{aligned}\oint |\mathfrak{f}_c(\boldsymbol{\alpha};\xi)|^{2t}\,\mathrm{d}\boldsymbol{\alpha} &\leqslant \oint |1+f(\boldsymbol{\alpha};X/p^c)|^{2t}\,\mathrm{d}\boldsymbol{\alpha}\\ &\ll_t 1+\oint |f(\boldsymbol{\alpha};X/p^c)|^{2t}\,\mathrm{d}\boldsymbol{\alpha}\\ &= 1+J_t(X/p^c).\end{aligned}$$

The desired conclusion follows on noting that diagonal solutions alone ensure that $J_t(X/M^c)\geqslant 1$. □

Our next preparatory manoeuvre concerns the initiation of the iterative procedure, and it is here that we fix our choice for p. It is convenient here and elsewhere to write $\mathbf{1}$ for the k-tuple $(1,\dots,1)$.

LEMMA 3.2. *There exists a prime number p with $M<p\leqslant 2M$ for which $J_{s+k}(X)\ll M^{2s}I_{0,1}(X)$.*

Proof. The quantity $J_{s+k}(X)$ counts the number of integral solutions of the system

$$\sum_{i=1}^{s+k}(x_i^j-y_i^j)=0\quad (1\leqslant j\leqslant k),$$

with $1\leqslant \mathbf{x},\mathbf{y}\leqslant X$. Let T_0 denote the number of such solutions in which $x_i=x_j$ for some i and j with $1\leqslant i<j\leqslant k$, and let T_1 denote the corresponding number of solutions with $x_i=x_j$ for no i and j with $1\leqslant i<j\leqslant k$.

On considering the underlying Diophantine system, one finds that

$$T_0\ll \oint f(2\boldsymbol{\alpha};X)f(\boldsymbol{\alpha};X)^{s+k-2}f(-\boldsymbol{\alpha};X)^{s+k}\,\mathrm{d}\boldsymbol{\alpha},$$

whence by Hölder's inequality, it follows that

$$T_0\ll \Big(\oint |f(\boldsymbol{\alpha};X)|^{2s+2k}\,\mathrm{d}\boldsymbol{\alpha}\Big)^{1-1/(s+k)}\Big(\oint |f(2\boldsymbol{\alpha};X)|^{2s+2k}\,\mathrm{d}\boldsymbol{\alpha}\Big)^{1/(2s+2k)}.$$

Thus, by a change of variables, we obtain the upper bound

$$T_0\ll (J_{s+k}(X))^{1-1/(2s+2k)}. \tag{3.14}$$

Consider next a solution $\mathbf{x},\mathbf{y}$ counted by T_1. Write

$$\Delta(\mathbf{x})=\prod_{1\leqslant i<j\leqslant k}|x_i-x_j|,$$

and note that $0<\Delta(\mathbf{x})<X^{k(k-1)}$. Let $\mathcal{P}$ denote any set of $[k^3/\theta]+1$ distinct prime numbers with $M<p\leqslant 2M$. Such a set exists by the Prime Number Theorem. It follows that

$$\prod_{p\in\mathcal{P}}p>M^{k^3/\theta}=X^{k^3}>\Delta(\mathbf{x}),$$

and hence at least one of the elements of $\mathcal{P}$ does not divide $\Delta(\mathbf{x})$. In particular, there exists a prime $p \in \mathcal{P}$ for which $x_i \equiv x_j \pmod{p}$ for no i and j with $1 \leqslant i < j \leqslant k$. On considering the underlying Diophantine system, we therefore see that

$$T_1 \ll \sum_{p \in \mathcal{P}} \oint \mathfrak{F}_0^{\mathbf{1}}(\boldsymbol{\alpha}; 0) f(\boldsymbol{\alpha}; X)^s f(-\boldsymbol{\alpha}; X)^{s+k} \, \mathrm{d}\boldsymbol{\alpha}.$$

Therefore, as a consequence of Schwarz's inequality, one finds that

$$\begin{aligned} T_1 &\ll \max_{p \in \mathcal{P}} \Big(\oint |\mathfrak{F}_0^{\mathbf{1}}(\boldsymbol{\alpha}; 0)^2 f(\boldsymbol{\alpha}; X)^{2s}| \, \mathrm{d}\boldsymbol{\alpha} \Big)^{1/2} \Big(\oint |f(\boldsymbol{\alpha}; X)|^{2s+2k} \, \mathrm{d}\boldsymbol{\alpha} \Big)^{1/2} \\ &= \max_{p \in \mathcal{P}} \Big(\oint |\mathfrak{F}_0^{\mathbf{1}}(\boldsymbol{\alpha}; 0)^2 \mathfrak{f}_0(\boldsymbol{\alpha}; 0)^{2s}| \, \mathrm{d}\boldsymbol{\alpha} \Big)^{1/2} \big(J_{s+k}(X) \big)^{1/2}. \end{aligned}$$

In this way, we deduce that a prime number p with $M < p \leqslant 2M$ exists for which

$$T_1 \ll (I_{0,0}(X))^{1/2} (J_{s+k}(X))^{1/2}. \tag{3.15}$$

On recalling that $J_{s+k}(X) = T_0 + T_1$, we find from (3.14) and (3.15) that

$$J_{s+k}(X) \ll 1 + I_{0,0}(X) \ll I_{0,0}(X). \tag{3.16}$$

Next, we split the summation in the definition (2.3) of $\mathfrak{f}_0(\boldsymbol{\alpha}; 0)$ into arithmetic progressions modulo p. Thus we obtain

$$\mathfrak{f}_0(\boldsymbol{\alpha}; 0) = \sum_{\xi=1}^{p} \mathfrak{f}_1(\boldsymbol{\alpha}; \xi),$$

whence by Hölder's inequality one has

$$|\mathfrak{f}_0(\boldsymbol{\alpha}; 0)|^{2s} \leqslant p^{2s-1} \sum_{\xi=1}^{p} |\mathfrak{f}_1(\boldsymbol{\alpha}; \xi)|^{2s}.$$

It therefore follows from (3.4) and (3.6) that

$$I_{0,0}(X) \ll M^{2s} \max_{1 \leqslant \xi \leqslant p} \max_{\boldsymbol{\sigma} \in \Sigma_k} \oint |\mathfrak{F}_0^{\boldsymbol{\sigma}}(\boldsymbol{\alpha}; 0)^2 \mathfrak{f}_1(\boldsymbol{\alpha}; \xi)^{2s}| \, \mathrm{d}\boldsymbol{\alpha} \leqslant M^{2s} I_{0,1}(X). \tag{3.17}$$

The conclusion of the lemma is obtained by substituting (3.17) into (3.16). □

We now fix the prime number p, once and for all, so that the upper bound $J_{s+k}(X) \ll M^{2s} I_{0,1}(X)$ holds.

4. **The auxiliary system of congruences**

The efficient congruencing process delivers a strong congruence condition on a subset of variables. In order to be useful in further congruencing activities, this condition must be converted into a restriction of certain variables to higher level arithmetic progressions. It is to this task that we attend in the present section.

When $\boldsymbol{\sigma} \in \Sigma_k$, denote by $\mathcal{B}_{a,b}^{\boldsymbol{\sigma}}(\mathbf{m};\xi,\eta)$ the set of solutions of the system of congruences

$$\sum_{i=1}^{k} \sigma_i (z_i - \eta)^j \equiv m_j \pmod{p^{jb}} \quad (1 \leqslant j \leqslant k), \tag{4.1}$$

with $1 \leqslant \mathbf{z} \leqslant p^{kb}$ and $\mathbf{z} \equiv \boldsymbol{\xi} \pmod{p^{a+1}}$ for some $\boldsymbol{\xi} \in \Xi_a(\xi)$.

LEMMA 4.1. *Suppose that a and b are nonnegative integers with $b > a$. Then*

$$\max_{1 \leqslant \xi \leqslant p^a} \max_{1 \leqslant \eta \leqslant p^b} \max_{\boldsymbol{\sigma} \in \Sigma_k} \mathit{card}\big(\mathcal{B}_{a,b}^{\boldsymbol{\sigma}}(\mathbf{m};\xi,\eta)\big) \leqslant k! p^{\frac{1}{2}k(k-1)(a+b)}.$$

Proof. Consider fixed integers a and b with $0 \leqslant a < b$, a fixed k-tuple $\boldsymbol{\sigma} \in \Sigma_k$, and fixed integers ξ and η with $1 \leqslant \xi \leqslant p^a$ and $1 \leqslant \eta \leqslant p^b$. Denote by $\mathcal{D}_1(\mathbf{n})$ the set of solutions of the system of congruences

$$\sum_{i=1}^{k} \sigma_i (z_i - \eta)^j \equiv n_j \pmod{p^{kb}} \quad (1 \leqslant j \leqslant k), \tag{4.2}$$

with $1 \leqslant \mathbf{z} \leqslant p^{kb}$ and $\mathbf{z} \equiv \boldsymbol{\xi} \pmod{p^{a+1}}$ for some $\boldsymbol{\xi} \in \Xi_a(\xi)$. Then it follows from (4.1) that we have

$$\operatorname{card}(\mathcal{B}_{a,b}^{\boldsymbol{\sigma}}(\mathbf{m};\xi,\eta)) = \sum_{\substack{n_1 \equiv m_1 \pmod{p^b} \\ 1 \leqslant n_1 \leqslant p^{kb}}} \cdots \sum_{\substack{n_k \equiv m_k \pmod{p^{kb}} \\ 1 \leqslant n_k \leqslant p^{kb}}} \operatorname{card}(\mathcal{D}_1(\mathbf{n})).$$

Counting the number of k-tuples $\mathbf{n}$ with $1 \leqslant \mathbf{n} \leqslant p^{kb}$ for which $n_j \equiv m_j \pmod{p^{jb}}$ $(1 \leqslant j \leqslant k)$, therefore, we see that

$$\operatorname{card}(\mathcal{B}_{a,b}^{\boldsymbol{\sigma}}(\mathbf{m};\xi,\eta)) \leqslant p^{\frac{1}{2}k(k-1)b} \max_{1 \leqslant \mathbf{n} \leqslant p^{kb}} \operatorname{card}(\mathcal{D}_1(\mathbf{n})). \tag{4.3}$$

We now examine the system (4.2). We begin by rewriting each variable z_i in the shape $z_i = p^a y_i + \xi$. In view of the hypothesis that $\mathbf{z} \equiv \boldsymbol{\xi} \pmod{p^{a+1}}$ for some $\boldsymbol{\xi} \in \Xi_a(\xi)$, we find that the k-tuple $\mathbf{y}$ satisfies the condition that $y_i \equiv y_j \pmod{p}$ for no i and j with $1 \leqslant i < j \leqslant k$. With this substitution in (4.2), we find by the Binomial Theorem that the set of solutions $\mathcal{D}_1(\mathbf{n})$ is in bijective correspondence with the set of solutions of the system of congruences

$$\sum_{l=0}^{j} \binom{j}{l} (\xi - \eta)^{j-l} p^{la} \sum_{i=1}^{k} \sigma_i y_i^l \equiv n_j \pmod{p^{kb}} \quad (1 \leqslant j \leqslant k), \tag{4.4}$$

with $1 \leqslant \mathbf{y} \leqslant p^{kb-a}$ and $y_i \equiv y_j \pmod{p}$ for no i and j with $1 \leqslant i < j \leqslant k$. Let $\mathbf{y} = \mathbf{w}$ be any solution of this system, if indeed a solution exists. Then it follows from (4.4) that all other solutions $\mathbf{y}$ satisfy the system of congruences

$$\sum_{l=0}^{j} \binom{j}{l} (\xi - \eta)^{j-l} p^{la} \sum_{i=1}^{k} \sigma_i (y_i^l - w_i^l) \equiv 0 \pmod{p^{kb}} \quad (1 \leqslant j \leqslant k). \tag{4.5}$$

By taking linear combinations of the congruences here, we find that the system (4.5) is equivalent to the new system

$$\sum_{i=1}^{k} \sigma_i y_i^j \equiv \sum_{i=1}^{k} \sigma_i w_i^j \pmod{p^{kb-ja}} \quad (1 \leqslant j \leqslant k).$$

Next, we write $\mathcal{D}_2(\mathbf{u})$ for the set of solutions of the system of congruences

$$\sum_{i=1}^{k} \sigma_i y_i^j \equiv u_j \pmod{p^{kb-ja}} \quad (1 \leqslant j \leqslant k),$$

with $1 \leqslant \mathbf{y} \leqslant p^{kb-a}$ and $y_i \equiv y_j \pmod{p}$ for no i and j with $1 \leqslant i < j \leqslant k$. Then it follows from our discussion thus far that

$$\operatorname{card}(\mathcal{D}_1(\mathbf{n})) \leqslant \max_{1 \leqslant \mathbf{u} \leqslant p^{kb-a}} \operatorname{card}(\mathcal{D}_2(\mathbf{u})). \tag{4.6}$$

Denote by $\mathcal{D}_3(\mathbf{v})$ the set of solutions of the system of congruences

$$\sum_{i=1}^{k} \sigma_i y_i^j \equiv v_j \pmod{p^{kb-a}} \quad (1 \leqslant j \leqslant k), \tag{4.7}$$

with $1 \leqslant \mathbf{y} \leqslant p^{kb-a}$ and $y_i \equiv y_j \pmod{p}$ for no i and j with $1 \leqslant i < j \leqslant k$. Then we have

$$\operatorname{card}(\mathcal{D}_2(\mathbf{u})) \leqslant \sum_{\substack{v_1 \equiv u_1 \pmod{p^{kb-a}} \\ 1 \leqslant v_1 \leqslant p^{kb-a}}} \dots \sum_{\substack{v_k \equiv u_k \pmod{p^{kb-ka}} \\ 1 \leqslant v_k \leqslant p^{kb-a}}} \operatorname{card}(\mathcal{D}_3(\mathbf{v})).$$

Counting the number of k-tuples $\mathbf{v}$ with $1 \leqslant \mathbf{v} \leqslant p^{kb-a}$ for which $v_j \equiv u_j \pmod{p^{kb-ja}}$ $(1 \leqslant j \leqslant k)$, therefore, we deduce that

$$\operatorname{card}(\mathcal{D}_2(\mathbf{u})) \leqslant p^{\frac{1}{2}k(k-1)a} \max_{1 \leqslant \mathbf{v} \leqslant p^{kb-a}} \operatorname{card}(\mathcal{D}_3(\mathbf{v})).$$

Consequently, in combination with (4.3) and (4.6), we have shown thus far that

$$\operatorname{card}(\mathcal{B}_{a,b}^{\boldsymbol{\sigma}}(\mathbf{m}; \xi, \eta)) \leqslant p^{\frac{1}{2}k(k-1)(a+b)} \max_{1 \leqslant \mathbf{v} \leqslant p^{kb-a}} \operatorname{card}(\mathcal{D}_3(\mathbf{v})). \tag{4.8}$$

Suppose now that $\mathbf{y} = \mathbf{z}$ is any solution of (4.7) belonging to $\mathcal{D}_3(\mathbf{v})$, if one exists. Then all other solutions $\mathbf{y}$ satisfy the system

$$\sum_{i=1}^{k} \sigma_i y_i^j \equiv \sum_{i=1}^{k} \sigma_i z_i^j \pmod{p^{kb-a}} \quad (1 \leqslant j \leqslant k).$$

Let $\mathcal{I}$ denote the set of indices i with $1 \leqslant i \leqslant k$ for which $\sigma_i = 1$, and let $\mathcal{J}$ denote the corresponding set of indices for which $\sigma_i = -1$. Then this system of congruences is equivalent to the new system

$$\sum_{i \in \mathcal{I}} y_i^j + \sum_{l \in \mathcal{J}} z_l^j \equiv \sum_{i \in \mathcal{I}} z_i^j + \sum_{l \in \mathcal{J}} y_l^j \pmod{p^{kb-a}} \quad (1 \leqslant j \leqslant k).$$

We are at liberty to assume that $p > k$. Consequently, from Newton's formulae relating the sums of powers of the roots of a polynomial with its coefficients, we find that

$$\prod_{i\in\mathcal{I}}(t-y_i)\prod_{l\in\mathcal{J}}(t-z_l)\equiv\prod_{j\in\mathcal{I}}(t-z_j)\prod_{m\in\mathcal{J}}(t-y_m)\pmod{p^{kb-a}}.$$

But $z_l \equiv z_m \pmod p$ for no l and m with $1 \leqslant l < m \leqslant k$. Then for each j with $j \in \mathcal{I}$, by putting $t = z_j$ we deduce that

$$\prod_{i\in\mathcal{I}}(z_j-y_i)\prod_{l\in\mathcal{J}}(z_j-z_l)\equiv 0\pmod{p^{kb-a}},$$

whence for some i with $i \in \mathcal{I}$, one has $y_i \equiv z_j \pmod{p^{kb-a}}$. Similarly, for each l with $l \in \mathcal{J}$, we deduce that for some m with $m \in \mathcal{J}$, one has $y_m \equiv z_l \pmod{p^{kb-a}}$. It follows that the sets $\{y_1, \dots, y_k\}$ and $\{z_1, \dots, z_k\}$ are mutually congruent modulo p^{kb-a}, whence $\operatorname{card}(\mathcal{D}_3(\mathbf{v})) \leqslant k!$. The conclusion of the lemma now follows at once from (4.8). □

5. **The conditioning process**

The mean value $I^{\boldsymbol{\sigma}}_{a,b}(X;\xi,\eta)$, defined via (3.4), is already in a form suitable for the extraction of an efficient congruence. Unfortunately, however, one would be poorly positioned to extract the next efficient congruence following the one at hand were one not to plan ahead by conditioning the auxiliary variables encoded by the exponential sum $\mathfrak{f}_b(\boldsymbol{\alpha};\eta)$. In this section we show that the factor $\mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{2s}$ occurring in (3.4) can, in essence, be replaced by the conditioned factor $\mathfrak{F}^{\boldsymbol{\tau}}_b(\boldsymbol{\alpha};\eta)^{2u}$. The latter involves k-tuples of variables in residue classes distinct modulo p^{b+1} and is suitable for subsequent congruencing operations.

LEMMA 5.1. *Let a and b be integers with $b > a \geqslant 0$. Then one has*

$$I_{a,b}(X) \ll K_{a,b}(X) + M^{k-1}I_{a,b+1}(X).$$

Proof. Consider fixed integers ξ and η with $1 \leqslant \xi \leqslant p^a$ and $1 \leqslant \eta \leqslant p^b$, and a k-tuple $\boldsymbol{\sigma} \in \Sigma_k$. Then on considering the underlying Diophantine system, one finds from (3.4) that $I^{\boldsymbol{\sigma}}_{a,b}(X;\xi,\eta)$ counts the number of integral solutions of the system

$$\sum_{i=1}^{k}\sigma_i(x_i^j-y_i^j)=\sum_{l=1}^{s}(v_l^j-w_l^j)\quad(1\leqslant j\leqslant k),\tag{5.1}$$

with

$$1\leqslant\mathbf{x},\mathbf{y},\mathbf{v},\mathbf{w}\leqslant X,\quad \mathbf{v}\equiv\mathbf{w}\equiv\eta\pmod{p^b}$$

and satisfying the property that there exist $\boldsymbol{\xi},\boldsymbol{\zeta}\in\Xi_a(\xi)$ for which

$$\mathbf{x}\equiv\boldsymbol{\xi}\pmod{p^{a+1}}\quad\text{and}\quad\mathbf{y}\equiv\boldsymbol{\zeta}\pmod{p^{a+1}}.$$

Let T_1 denote the number of integral solutions $\mathbf{x}, \mathbf{y}, \mathbf{v}, \mathbf{w}$ of the system (5.1), counted by $I_{a,b}^{\boldsymbol{\sigma}}(X;\xi,\eta)$, in which the $2s$ integers $v_1,\dots,v_s$ and $w_1,\dots,w_s$ together lie in at most $k-1$ distinct residue classes modulo p^{b+1}, and let T_2 denote the corresponding number of solutions in which the $2s$ integers $v_1,\dots,v_s$ and $w_1,\dots,w_s$ together occupy at least k distinct residue classes modulo p^{b+1}. Then we have

$$I_{a,b}^{\boldsymbol{\sigma}}(X;\xi,\eta) \leqslant T_1 + T_2.$$

On considering the underlying Diophantine system, it is apparent that

$$T_1 \ll \sum_{\substack{1\leqslant \eta_1,\dots,\eta_{k-1}\leqslant p^{b+1}\\ \boldsymbol{\eta}\equiv\eta \pmod{p^b}}} \sum_{\mathbf{0}\leqslant\mathbf{e}\leqslant 2s} \oint |\mathfrak{F}_a^{\boldsymbol{\sigma}}(\boldsymbol{\alpha};\xi)^2 \mathfrak{f}_{b+1}(\boldsymbol{\alpha};\eta_1)^{e_1}\dots\mathfrak{f}_{b+1}(\boldsymbol{\alpha};\eta_{k-1})^{e_{k-1}}|\,\mathrm{d}\boldsymbol{\alpha},$$

in which the summation over $\mathbf{e}$ is subject to the condition

$$e_1+e_2+\cdots+e_{k-1}=2s.$$

In view of the elementary inequality

$$|z_1\cdots z_n| \leqslant |z_1|^n+\cdots+|z_n|^n,$$

we find that

$$|\mathfrak{f}_{b+1}(\boldsymbol{\alpha};\eta_1)^{e_1}\cdots\mathfrak{f}_{b+1}(\boldsymbol{\alpha};\eta_{k-1})^{e_{k-1}}| \leqslant \sum_{i=1}^{k-1}|\mathfrak{f}_{b+1}(\boldsymbol{\alpha};\eta_i)|^{2s}.$$

Thus we deduce that

$$\begin{aligned} T_1 &\ll \sum_{\substack{1\leqslant \eta_1,\dots,\eta_{k-1}\leqslant p^{b+1}\\ \boldsymbol{\eta}\equiv\eta \pmod{p^b}}} \sum_{i=1}^{k-1} \oint |\mathfrak{F}_a^{\boldsymbol{\sigma}}(\boldsymbol{\alpha};\xi)^2 \mathfrak{f}_{b+1}(\boldsymbol{\alpha};\eta_i)^{2s}|\,\mathrm{d}\boldsymbol{\alpha} \\ &\ll p^{k-1} \max_{1\leqslant\eta_0\leqslant p^{b+1}} I_{a,b+1}^{\boldsymbol{\sigma}}(X;\xi,\eta_0). \end{aligned} \tag{5.2}$$

We turn our attention next to the solutions $\mathbf{x}, \mathbf{y}, \mathbf{v}, \mathbf{w}$ counted by T_2. The $2s$ integers $v_1,\dots,v_s$ and $w_1,\dots,w_s$ now together contain at least k distinct residue classes modulo p^{b+1}. By relabelling variables if necessary, therefore, there is no loss of generality in supposing that $v_1,\dots,v_k$ lie in distinct residue classes modulo p^{b+1}. We emphasise here that, not only the indices of the elements v_i may need adjustment in this relabelling process, but also certain elements w_j may need to be renamed as one of the new integers $v_1,\dots,v_k$. In the former cases, the associated signs indexed by τ_i will be $+1$, and in the latter cases they will be -1. On considering the underlying Diophantine system, we thus deduce that for some $\boldsymbol{\tau}\in\Sigma_k$, one has

$$T_2 \ll \oint |\mathfrak{F}_a^{\boldsymbol{\sigma}}(\boldsymbol{\alpha};\xi)|^2 \mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)\mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{s-r_+}\mathfrak{f}_b(-\boldsymbol{\alpha};\eta)^{s-r_-}\,\mathrm{d}\boldsymbol{\alpha}.$$

Here, we have written r_+ for the number of the coordinates of $\boldsymbol{\tau}$ that are $+1$ and r_- for the number that are -1. Thus, in particular, one has $r_+ + r_- = k$. On recalling that $s = uk$, an application of Hölder's inequality leads from here to the bound

$$T_2 \ll \Big(\oint |\mathfrak{F}_a^{\boldsymbol{\sigma}}(\boldsymbol{\alpha};\xi)^2 \mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)^{2u}| \, \mathrm{d}\boldsymbol{\alpha} \Big)^{1/(2u)} \Big(\oint |\mathfrak{F}_a^{\boldsymbol{\sigma}}(\boldsymbol{\alpha};\xi)^2 \mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{2s}| \, \mathrm{d}\boldsymbol{\alpha} \Big)^{1-1/(2u)}.$$

Hence, in view of the definitions (3.4) and (3.5), we arrive at the estimate

$$T_2 \ll (K_{a,b}^{\boldsymbol{\sigma},\boldsymbol{\tau}}(X;\xi,\eta))^{1/(2u)} (I_{a,b}^{\boldsymbol{\sigma}}(X;\xi,\eta))^{1-1/(2u)}. \tag{5.3}$$

Combining (5.2) and (5.3), and recalling (3.6) and (3.7), we deduce that

$$I_{a,b}(X) \ll M^{k-1} I_{a,b+1}(X) + (K_{a,b}(X))^{1/(2u)} (I_{a,b}(X))^{1-1/(2u)}.$$

The conclusion of the lemma now follows on disentangling this inequality. □

Repeated application of Lemma 5.1 shows that whenever a, b and H are nonnegative integers with $b > a \geqslant 0$, then

$$I_{a,b}(X) \ll \sum_{h=0}^{H-1} M^{h(k-1)} K_{a,b+h}(X) + M^{H(k-1)} I_{a,b+H}(X). \tag{5.4}$$

Since for large values of H, quantities of the type $I_{a,b+H}(X)$ are an irritant to our argument, we show in the next lemma that values of H exceeding $\frac{1}{2}(b-a)$ are harmless.

Lemma 5.2. *Let a, b and H be nonnegative integers with*

$$0 < \tfrac{1}{2}(b-a) \leqslant H \leqslant \theta^{-1} - b.$$

Then one has

$$M^{H(k-1)} I_{a,b+H}(X) \ll M^{-H/2} (X/M^b)^{2s} (X/M^a)^{2k - \frac{1}{2}k(k+1) + \eta_{s+k}}.$$

Proof. On considering the underlying Diophantine systems, it follows from (3.3) and (3.4) that when $1 \leqslant \xi \leqslant p^a$ and $1 \leqslant \eta \leqslant p^{b+H}$, and $\boldsymbol{\sigma} \in \Sigma_k$, one has

$$I_{a,b+H}^{\boldsymbol{\sigma}}(X;\xi,\eta) \leqslant \oint |\mathfrak{f}_a(\boldsymbol{\alpha};\xi)^{2k} \mathfrak{f}_{b+H}(\boldsymbol{\alpha};\eta)^{2s}| \, \mathrm{d}\boldsymbol{\alpha}.$$

Then an application of Hölder's inequality in combination with Lemma 3.1 leads to the upper bound

$$\begin{aligned} I_{a,b+H}^{\boldsymbol{\sigma}}(X;\xi,\eta) &\leqslant \Big(\oint |\mathfrak{f}_a(\boldsymbol{\alpha};\xi)|^{2s+2k} \, \mathrm{d}\boldsymbol{\alpha} \Big)^{k/(s+k)} \Big(\oint |\mathfrak{f}_{b+H}(\boldsymbol{\alpha};\eta)|^{2s+2k} \, \mathrm{d}\boldsymbol{\alpha} \Big)^{s/(s+k)} \\ &\ll (J_{s+k}(X/M^a))^{k/(s+k)} (J_{s+k}(X/M^{b+H}))^{s/(s+k)}. \end{aligned}$$

Consequently, in view of (3.2), we have

$$\begin{aligned} I_{a,b+H}(X) &\ll ((X/M^a)^{k/(s+k)} (X/M^{b+H})^{s/(s+k)})^{2s+2k-\frac{1}{2}k(k+1)+\eta_{s+k}+\delta} \\ &\ll X^{\delta} (X/M^a)^{2k-\frac{1}{2}k(k+1)+\eta_{s+k}} (X/M^b)^{2s} \Upsilon, \end{aligned} \tag{5.5}$$

where

$$\Upsilon = (M^{b-a+H})^{\frac{1}{2}k(k+1)s/(s+k)}M^{-2sH}.$$

But when $s \geqslant k^2$ and $H \geqslant \frac{1}{2}(b-a)$, one has

$$\begin{aligned}\frac{s}{s+k}\left(2(s+k)-\tfrac{1}{2}k(k+1)\right)H &\geqslant \frac{s}{s+k}\left(\tfrac{3}{2}k(k+1)\right)H\\ &\geqslant \frac{s}{s+k}\left(\tfrac{1}{2}k(k+1)\right)(b-a)+\tfrac{1}{2}k^2H.\end{aligned}$$

Thus we see that for $k \geqslant 2$, one has

$$M^{H(k-1)}\Upsilon \leqslant M^{H(k-1-\frac{1}{2}k^2)} \leqslant M^{-H},$$

whence

$$X^{\delta}M^{H(k-1)}\Upsilon \leqslant M^{-H/2}.$$

The conclusion of the lemma follows on substituting this estimate into (5.5). □

Combining Lemma 5.2 with the upper bound (5.4), we may conclude as follows. Here, as usual, when $\beta \in \mathbb{R}$, we write $\lceil \beta \rceil$ for the least integer no smaller than β.

LEMMA 5.3. *Let a and b be integers with $0 \leqslant a < b$, and put $H = \lceil \frac{1}{2}(b-a) \rceil$. Suppose that $b + H \leqslant \theta^{-1}$. Then there exists an integer h with $0 \leqslant h < H$ having the property that*

$$I_{a,b}(X) \ll M^{h(k-1)}K_{a,b+h}(X) + M^{-H/2}(X/M^b)^{2s}(X/M^a)^{2k-\frac{1}{2}k(k+1)+\eta_{s+k}}.$$

By making use of the special case of Lemma 5.3 in which $a = 0$ and $b = 1$, we are able to refine Lemma 3.2 into a form more directly applicable.

LEMMA 5.4. *One has $J_{s+k}(X) \ll M^{2s}K_{0,1}(X)$.*

Proof. Observe first that when $a = 0$ and $b = 1$, then $\lceil \frac{1}{2}(b-a) \rceil = 1$. Thus we deduce from Lemma 5.3 that

$$I_{0,1}(X) \ll K_{0,1}(X) + M^{-1/2}(X/M)^{2s}X^{2k-\frac{1}{2}k(k+1)+\eta_{s+k}}.$$

Since we may suppose that $M^{1/2} > X^{4\delta}$, it follows from Lemma 3.2 that

$$J_{s+k}(X) \ll M^{2s}I_{0,1}(X) \ll M^{2s}K_{0,1}(X) + X^{2s+2k-\frac{1}{2}k(k+1)+\eta_{s+k}-2\delta}.$$

But in view of (3.10), we have

$$J_{s+k}(X) \gg X^{2s+2k-\frac{1}{2}k(k+1)+\eta_{s+k}-\delta},$$

and hence we arrive at the upper bound

$$J_{s+k}(X) \ll M^{2s}K_{0,1}(X) + X^{-\delta}J_{s+k}(X).$$

The conclusion of the lemma follows on disentangling this inequality. □

6. **The efficient congruencing step**

The mean value $K_{a,b}(X)$ contains a powerful latent congruence condition. Our task in this section is to convert this condition into one that may be exploited by means of an iterative procedure.

LEMMA 6.1. *Suppose that a and b are integers with $0 \leqslant a < b \leqslant \theta^{-1}$. Then one has*

$$K_{a,b}(X) \ll M^{\frac{1}{2}k(k-1)(b+a)}(M^{kb-a})^k \left(J_{s+k}(X/M^b)\right)^{1-k/s} \left(I_{b,kb}(X)\right)^{k/s}.$$

Proof. Consider fixed integers ξ and η with $1 \leqslant \xi \leqslant p^a$ and $1 \leqslant \eta \leqslant p^b$, and k-tuples $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \Sigma_k$. Then on considering the underlying Diophantine system, one finds from (3.5) that $K_{a,b}^{\boldsymbol{\sigma},\boldsymbol{\tau}}(X;\xi,\eta)$ counts the number of integral solutions of the system

$$\sum_{i=1}^{k} \sigma_i(x_i^j - y_i^j) = \sum_{l=1}^{u}\sum_{m=1}^{k} \tau_m(v_{lm}^j - w_{lm}^j) \quad (1 \leqslant j \leqslant k), \tag{6.1}$$

in which, for some $\boldsymbol{\xi}, \boldsymbol{\zeta} \in \Xi_a(\xi)$, one has

$$1 \leqslant \mathbf{x}, \mathbf{y} \leqslant X, \quad \mathbf{x} \equiv \boldsymbol{\xi} \pmod{p^{a+1}} \quad \text{and} \quad \mathbf{y} \equiv \boldsymbol{\zeta} \pmod{p^{a+1}},$$

and for $1 \leqslant l \leqslant u$, for some $\boldsymbol{\eta}_l, \boldsymbol{\nu}_l \in \Xi_b(\eta)$, one has

$$1 \leqslant \mathbf{v}_l, \mathbf{w}_l \leqslant X, \quad \mathbf{v}_l \equiv \boldsymbol{\eta}_l \pmod{p^{b+1}} \quad \text{and} \quad \mathbf{w}_l \equiv \boldsymbol{\nu}_l \pmod{p^{b+1}}.$$

By applying the Binomial Theorem, we see that the system (6.1) is equivalent to the new system of equations

$$\sum_{i=1}^{k} \sigma_i((x_i-\eta)^j - (y_i-\eta)^j) = \sum_{l=1}^{u}\sum_{m=1}^{k} \tau_m((v_{lm}-\eta)^j - (w_{lm}-\eta)^j) \quad (1 \leqslant j \leqslant k). \tag{6.2}$$

But in any solution $\mathbf{x}, \mathbf{y}, \mathbf{v}, \mathbf{w}$ counted by $K_{a,b}^{\boldsymbol{\sigma},\boldsymbol{\tau}}(X;\xi,\eta)$, one has $\mathbf{v} \equiv \mathbf{w} \equiv \eta \pmod{p^b}$. We therefore deduce from (6.2) that

$$\sum_{i=1}^{k} \sigma_i(x_i - \eta)^j \equiv \sum_{i=1}^{k} \sigma_i(y_i - \eta)^j \pmod{p^{jb}} \quad (1 \leqslant j \leqslant k). \tag{6.3}$$

Recall the notation from the preamble to Lemma 4.1, and write

$$\mathfrak{G}_{a,b}^{\boldsymbol{\sigma}}(\boldsymbol{\alpha};\xi,\eta;\mathbf{m}) = \sum_{\boldsymbol{\zeta} \in \mathcal{B}_{a,b}^{\boldsymbol{\sigma}}(\mathbf{m};\xi,\eta)} \prod_{i=1}^{k} \mathfrak{f}_{kb}(\sigma_i\boldsymbol{\alpha};\zeta_i).$$

Then on considering the underlying Diophantine system, it follows from (6.1) and (6.3) that

$$K_{a,b}^{\boldsymbol{\sigma},\boldsymbol{\tau}}(X;\xi,\eta) = \sum_{m_1=1}^{p^b} \cdots \sum_{m_k=1}^{p^{kb}} \oint |\mathfrak{G}_{a,b}^{\boldsymbol{\sigma}}(\boldsymbol{\alpha};\xi,\eta;\mathbf{m})^2 \mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)^{2u}| \, \mathrm{d}\boldsymbol{\alpha}. \tag{6.4}$$

An application of Cauchy's inequality in combination with Lemma 4.1 yields the upper bound

$$|\mathfrak{G}_{a,b}^{\boldsymbol{\sigma}}(\boldsymbol{\alpha};\xi,\eta;\mathbf{m})|^2 \leqslant \operatorname{card}(\mathcal{B}_{a,b}^{\boldsymbol{\sigma}}(\mathbf{m};\xi,\eta)) \sum_{\boldsymbol{\zeta}\in\mathcal{B}_{a,b}^{\boldsymbol{\sigma}}(\mathbf{m};\xi,\eta)} \prod_{i=1}^{k} |\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta_i)|^2 \ll M^{\frac{1}{2}k(k-1)(a+b)} \sum_{\boldsymbol{\zeta}\in\mathcal{B}_{a,b}^{\boldsymbol{\sigma}}(\mathbf{m};\xi,\eta)} \prod_{i=1}^{k} |\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta_i)|^2. \tag{6.5}$$

Next, on substituting (6.5) into (6.4) and considering the underlying Diophantine system, we deduce that

$$K_{a,b}^{\boldsymbol{\sigma},\boldsymbol{\tau}}(X;\xi,\eta) \ll M^{\frac{1}{2}k(k-1)(a+b)} \sum_{\substack{1\leqslant\boldsymbol{\zeta}\leqslant p^{kb}\\ \boldsymbol{\zeta}\equiv\xi\ (\mathrm{mod}\ p^a)}} \oint \Big(\prod_{i=1}^{k} |\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta_i)|^2\Big) |\mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)|^{2u}\,\mathrm{d}\boldsymbol{\alpha}. \tag{6.6}$$

Notice that the utility of the conditioning of the two initial blocks of k variables in (6.1) has now expired, and indeed in (6.6) this conditioning is abandoned without ill consequences for the argument to follow.

Observe next that by Hölder's inequality, one has

$$\sum_{\substack{1\leqslant\boldsymbol{\zeta}\leqslant p^{kb}\\ \boldsymbol{\zeta}\equiv\xi\ (\mathrm{mod}\ p^a)}} \prod_{i=1}^{k} |\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta_i)|^2 = \Big(\sum_{\substack{1\leqslant\zeta\leqslant p^{kb}\\ \zeta\equiv\xi\ (\mathrm{mod}\ p^a)}} |\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta)|^2\Big)^k \leqslant (p^{kb-a})^{k-1} \sum_{\substack{1\leqslant\zeta\leqslant p^{kb}\\ \zeta\equiv\xi\ (\mathrm{mod}\ p^a)}} |\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta)|^{2k}.$$

Then it follows from (6.6) that

$$K_{a,b}^{\boldsymbol{\sigma},\boldsymbol{\tau}}(X;\xi,\eta) \ll M^{\frac{1}{2}k(k-1)(a+b)}(M^{kb-a})^k \max_{1\leqslant\zeta\leqslant p^{kb}} \oint |\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta)^{2k}\mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)^{2u}|\,\mathrm{d}\boldsymbol{\alpha}. \tag{6.7}$$

On recalling that $s = uk$, an application of Hölder's inequality supplies the bound

$$\oint |\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta)^{2k}\mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)^{2u}|\,\mathrm{d}\boldsymbol{\alpha} \leqslant U_1^{1-k/s}U_2^{k/s}, \tag{6.8}$$

where

$$U_1 = \oint |\mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)|^{2u+2}\,\mathrm{d}\boldsymbol{\alpha}$$

and

$$U_2 = \oint |\mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)^2\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta)^{2s}|\,\mathrm{d}\boldsymbol{\alpha}.$$

On considering the underlying Diophantine system, it follows from Lemma 3.1 that

$$U_1 \leqslant \oint |\mathfrak{f}_b(\boldsymbol{\alpha};\eta)|^{2s+2k}\,\mathrm{d}\boldsymbol{\alpha} \ll J_{s+k}(X/M^b).$$

Thus, on recalling the definition (3.4), we find that

$$\begin{aligned}\oint |\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta)^{2k}\mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)^{2u}|\,\mathrm{d}\boldsymbol{\alpha} &\ll (J_{s+k}(X/M^b))^{1-k/s}(I^{\boldsymbol{\tau}}_{b,kb}(X;\eta,\zeta))^{k/s}\\ &\ll (J_{s+k}(X/M^b))^{1-k/s}(I_{b,kb}(X))^{k/s}.\end{aligned}$$

Finally, on substituting the latter estimate into (6.7), the conclusion of the lemma is immediate. □

Before proceeding further, we pause to extract a crude but simple bound for $K_{a,b}(X)$ of value when b is large.

Lemma 6.2. *Suppose that a and b are integers with $0 \leqslant a < b \leqslant \theta^{-1}$. Then*

$$[[K_{a,b}(X)]] \ll X^{\eta_{s+k}+\delta}(M^{b-a})^{\frac{1}{2}k(k+1)}.$$

Proof. Consider fixed integers ξ and η with $1 \leqslant \xi \leqslant p^a$ and $1 \leqslant \eta \leqslant p^b$, and k-tuples $\boldsymbol{\sigma},\boldsymbol{\tau} \in \Sigma_k$. On considering the underlying Diophantine system and applying Hölder's inequality, we deduce from (3.5) that

$$\begin{aligned}K^{\boldsymbol{\sigma},\boldsymbol{\tau}}_{a,b}(X;\xi,\eta) &\leqslant \oint |\mathfrak{f}_a(\boldsymbol{\alpha};\xi)^{2k}\mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{2s}|\,\mathrm{d}\boldsymbol{\alpha}\\ &\leqslant \Big(\oint |\mathfrak{f}_a(\boldsymbol{\alpha};\xi)|^{2s+2k}\,\mathrm{d}\boldsymbol{\alpha}\Big)^{k/(s+k)}\Big(\oint |\mathfrak{f}_b(\boldsymbol{\alpha};\eta)|^{2s+2k}\,\mathrm{d}\boldsymbol{\alpha}\Big)^{s/(s+k)}.\end{aligned}$$

In view of the hypothesis $b \leqslant \theta^{-1}$, we therefore deduce from Lemma 3.1 that

$$K_{a,b}(X) \ll (J_{s+k}(X/M^a))^{k/(s+k)}(J_{s+k}(X/M^b))^{s/(s+k)}.$$

Consequently, on recalling (3.9) and (3.11), it follows that

$$\begin{aligned}[[K_{a,b}(X)]] &\ll \frac{X^\delta\left((X/M^a)^{k/(s+k)}(X/M^b)^{s/(s+k)}\right)^{2s+2k-\frac{1}{2}k(k+1)+\eta_{s+k}}}{(X/M^b)^{2s}(X/M^a)^{2k-\frac{1}{2}k(k+1)}}\\ &\ll X^{\eta_{s+k}+\delta}(M^{b-a})^{\frac{1}{2}k(k+1)s/(s+k)}.\end{aligned}$$

The conclusion of the lemma is now immediate. □

By substituting the estimate supplied by Lemma 5.3 into the conclusion of Lemma 6.1, we obtain the basic iterative relation.

Lemma 6.3. *Suppose that a and b are integers with $0 \leqslant a < b \leqslant \frac{2}{3}(k\theta)^{-1}$. Put $H = \lceil \frac{1}{2}(k-1)b \rceil$. Then there exists an integer h, with $0 \leqslant h < H$, having*

the property that

$$[[K_{a,b}(X)]] \ll X^{\delta}M^{-7kh/4}(X/M^b)^{\eta_{s+k}(1-k/s)}[[K_{b,kb+h}(X)]]^{k/s} + M^{-kH/(3s)}(X/M^b)^{\eta_{s+k}}.$$

Proof. On recalling (3.9), it follows from Lemma 6.1 that

(6.9) $$[[K_{a,b}(X)]] \ll (M^b)^{2s}(M^a)^{2k-\frac{1}{2}k(k+1)}M^{\frac{1}{2}k(k-1)(b+a)}(M^{kb-a})^kT_1^{1-k/s}T_2^{k/s},$$

where

$$T_1 = \frac{J_{s+k}(X/M^b)}{X^{2s+2k-\frac{1}{2}k(k+1)}} \quad \text{and} \quad T_2 = \frac{I_{b,kb}(X)}{X^{2s+2k-\frac{1}{2}k(k+1)}}.$$

But in view of (3.2), one has

$$T_1 \ll (M^{-b})^{2s+2k-\frac{1}{2}k(k+1)}(X/M^b)^{\eta_{s+k}+\delta}. \tag{6.10}$$

Write $H = \lceil \frac{1}{2}(k-1)b \rceil$, and note that the hypotheses of the statement of the lemma ensure that

$$kb + H \leqslant kb + \tfrac{1}{2}(k-1)b + \tfrac{1}{2} \leqslant \tfrac{3}{2}kb \leqslant \theta^{-1}.$$

Consequently, it follows from Lemma 5.3 that there exists an integer h with $0 \leqslant h < H$ having the property that

$$T_2 \ll \frac{M^{h(k-1)}K_{b,kb+h}(X)}{X^{2s+2k-\frac{1}{2}k(k+1)}} + \frac{M^{-H/2}(X/M^b)^{\eta_{s+k}}}{(M^{kb})^{2s}(M^b)^{2k-\frac{1}{2}k(k+1)}}.$$

On recalling (3.9), we therefore see that

$$T_2 \ll (M^{-kb})^{2s}(M^{-b})^{2k-\frac{1}{2}k(k+1)}\Omega, \tag{6.11}$$

in which we have written

$$\Omega = M^{-(2s-k+1)h}[[K_{b,kb+h}(X)]] + M^{-H/2}(X/M^b)^{\eta_{s+k}}.$$

Substituting (6.10) and (6.11) into (6.9), we deduce that

$$[[K_{a,b}(X)]] \ll M^{\omega(a,b)}(X/M^b)^{(1-k/s)(\eta_{s+k}+\delta)}\Omega^{k/s},$$

in which we have written

$$\begin{aligned}\omega(a,b) =& 2sb + (2k - \tfrac{1}{2}k(k+1))a + \tfrac{1}{2}k(k-1)(b+a) + k(kb-a)\\ &- (1-k/s)(2s + 2k - \tfrac{1}{2}k(k+1))b - (2skb + (2k - \tfrac{1}{2}k(k+1))b)k/s.\end{aligned}$$

A modicum of computation reveals that $\omega(a,b) = 0$, and thus we may infer that

$$[[K_{a,b}(X)]] \ll (M^{-H/2})^{k/s}(X/M^b)^{\eta_{s+k}+\delta(1-k/s)} + X^{\delta}M^{-(2s-k+1)hk/s}(X/M^b)^{\eta_{s+k}(1-k/s)}[[K_{b,kb+h}(X)]]^{k/s}.$$

The conclusion of the lemma follows on noting that δ may be assumed small enough that $(X/M^b)^{\delta(1-k/s)} \ll M^{kH/(6s)}$, and further that the assumptions $s \geqslant k^2$ and $k \geqslant 2$ together imply that $2s-k+1 \geqslant \frac{7}{4}s$. $\square$

7. The iterative process

The estimate supplied by Lemma 5.4 bounds $J_{s+k}(X)$ in terms of $K_{0,1}(X)$, and Lemma 6.3 relates $K_{a,b}(X)$, for $b > a \geqslant 0$, to $K_{b,kb+h}(X)$ for some integer h with $0 \leqslant h \leqslant \frac{1}{2}(k-1)b$. By repeatedly applying Lemma 6.3, therefore, we are able to bound $J_{s+k}(X)$ in terms of the quantity $K_{c,d}(X)$, with c and d essentially as large as we please. Unfortunately, this process is not particularly simple to control, largely owing to the possibility that at any point in our iteration, a value of h in the expression $K_{b,kb+h}(X)$ may be forced upon us with $h > 0$. This defect in our procedure may accelerate us too rapidly towards the final step of the iteration. Our goal in this section, therefore, is to control the iterative process at a fine enough level that its potential is not substantially eroded.

LEMMA 7.1. *Suppose that a and b are integers with $0 \leqslant a < b \leqslant \frac{2}{3}(k\theta)^{-1}$. Suppose, in addition, that there exist nonnegative numbers ψ, c and γ, with $c \leqslant (2s/k)^N$, for which*

$$X^{\eta_{s+k}(1+\psi\theta)} \ll X^{c\delta} M^{-\gamma}[[K_{a,b}(X)]]. \tag{7.1}$$

Then, for some nonnegative integer h with $h \leqslant \frac{1}{2}(k-1)b$, one has

$$X^{\eta_{s+k}(1+\psi'\theta)} \ll X^{c'\delta} M^{-\gamma'}[[K_{a',b'}(X)]],$$

where

$$\psi' = (s/k)\psi + (s/k-1)b, \quad c' = (s/k)(c+1), \quad \gamma' = (s/k)\gamma + \tfrac{7}{4}sh,$$
$$a' = b \quad \text{and} \quad b' = kb+h.$$

Proof. Since we may suppose that $c \leqslant (2s/k)^N$ and $\delta < (Ns)^{-3N}$, we have

$$c\delta < s^{-2N}/3 < \theta/(3s),$$

and hence $X^{c\delta} < M^{1/(3s)}$. In addition, one has $M^{1/(3s)} > X^{\delta}$. Consequently, it follows from Lemma 6.3 that there exists an integer h with $0 \leqslant h < \lceil \frac{1}{2}(k-1)b \rceil$ with the property that

$$[[K_{a,b}(X)]] \ll M^{-k/(3s)} X^{\eta_{s+k}} + X^{\delta} M^{-7kh/4} (X/M^b)^{(1-k/s)\eta_{s+k}} [[K_{b,kb+h}(X)]]^{k/s}.$$

In view of the hypothesised upper bound (7.1), therefore, we deduce that

$$X^{\eta_{s+k}(1+\psi\theta)} \ll X^{\eta_{s+k}-\delta} + X^{(c+1)\delta} M^{-\gamma-7kh/4} (X/M^b)^{(1-k/s)\eta_{s+k}} [[K_{b,kb+h}(X)]]^{k/s},$$

whence

$$X^{\eta_{s+k}(k/s+(\psi+(1-k/s)b)\theta)} \ll X^{(c+1)\delta} M^{-\gamma-7kh/4} [[K_{b,kb+h}(X)]]^{k/s}.$$

The conclusion of the lemma follows on raising left- and right-hand sides here to the power s/k. □

Repeated application of Lemma 7.1 provides a series of upper bounds for η_{s+k}. What remains is to ensure that the upper bound $b \leqslant \frac{2}{3}(k\theta)^{-1}$, required by the hypotheses of the lemma, does not preclude the possibility of making many iterations.

LEMMA 7.2. *Whenever $s \geqslant k^2$, one has $\eta_{s+k} = 0$.*

Proof. In the final moments of our proof, we find it convenient to restrict s to be k^2. However, our argument is made more illuminating by avoiding this restriction in the opening stages. We may suppose that $\eta_{s+k} > 0$, for otherwise there is nothing to prove. We begin by defining three sequences (a_n), (b_n), (h_n) of nonnegative integers for $0 \leqslant n \leqslant N$. We put $a_0 = 0$ and $b_0 = 1$. Then, when $0 \leqslant n < N$, we fix any integer h_n with $0 \leqslant h_n \leqslant \frac{1}{2}(k-1)b_n$, and then define

$$a_{n+1} = b_n \quad \text{and} \quad b_{n+1} = kb_n + h_n. \tag{7.2}$$

Next we define the auxiliary sequences (ψ_n), (c_n), (γ_n) of nonnegative real numbers for $0 \leqslant n \leqslant N$ by putting $\psi_0 = 0$, $c_0 = 1$, $\gamma_0 = 0$. Then, for $0 \leqslant n < N$, we define

$$\psi_{n+1} = (s/k)\psi_n + (s/k - 1)b_n, \tag{7.3}$$

$$c_{n+1} = (s/k)(c_n + 1), \tag{7.4}$$

$$\gamma_{n+1} = (s/k)\gamma_n + \tfrac{7}{4}sh_n. \tag{7.5}$$

Notice here that an inductive argument readily confirms that for $0 \leqslant n \leqslant N$, one has

$$c_n = \frac{2s-k}{s-k}\Big(\frac{s}{k}\Big)^n - \frac{s}{s-k} \leqslant \Big(2 + \frac{1}{k-1}\Big)\Big(\frac{s}{k}\Big)^n \leqslant 3(s/k)^n.$$

We claim that a choice may be made for the sequence (h_n) in such a manner that for $0 \leqslant n \leqslant N$, one has

$$b_n < 2(s/k)^n \tag{7.6}$$

and

$$X^{\eta_{s+k}(1+\psi_n\theta)} \ll X^{c_n\delta}M^{-\gamma_n}[[K_{a_n,b_n}(X)]]. \tag{7.7}$$

When $n = 0$, the validity of the relation (7.6) follows by definition, whilst (7.7) is immediate from (3.9), (3.10) and Lemma 5.4, since the latter together imply that

$$X^{\eta_{s+k}-\delta} < [[J_{s+k}(X)]] \ll [[K_{0,1}(X)]].$$

We prepare the ground for the treatment of larger indices n with a preliminary discussion of the recurrence relations (7.2) to (7.5). Observe first that when $m \geqslant 0$, one has

$$\gamma_{m+1} - \tfrac{7}{4}k^2 b_{m+1} \geqslant \gamma_{m+1} - \tfrac{7}{4}s b_{m+1} = (s/k)(\gamma_m - \tfrac{7}{4}k^2 b_m).$$

But $\gamma_0 - \frac{7}{4}k^2 b_0 = -\frac{7}{4}k^2$, and so it follows by induction that when $0 \leqslant m \leqslant N$, one has

$$\gamma_m \geqslant \tfrac{7}{4}k^2(b_m - (s/k)^m). \tag{7.8}$$

Suppose now that the desired conclusions (7.6) and (7.7) have been established for the index $n < N$. Then as a consequence of (7.6) one has $kb_n\theta < k(s/k)^{n-N-1} < \frac{2}{3}$, whence $b_n < \frac{2}{3}(k\theta)^{-1}$. We may therefore apply Lemma 7.1 to deduce from (7.7) that there exists a nonnegative integer h, with $h \leqslant \frac{1}{2}(k-1)b_n$, for which one has the upper bound

$$X^{\eta_{s+k}(1+\psi'\theta)} \ll X^{c'\delta} M^{-\gamma'} [[K_{a',b'}(X)]], \tag{7.9}$$

where

$$a' = b_n = a_{n+1}, \quad b' = kb_n + h, \tag{7.10}$$

$$\psi' = (s/k)\psi_n + (s/k - 1)b_n = \psi_{n+1},$$

$$c' = (s/k)(c_n + 1) = c_{n+1},$$

$$\gamma' = (s/k)\gamma_n + \tfrac{7}{4}sh. \tag{7.11}$$

Let us suppose, if possible, that $b' \geqslant 2(s/k)^{n+1}$. The relations (7.10) and (7.11) then combine with (7.8) to show that

$$\begin{aligned} \gamma' &= (s/k)\gamma_n + \tfrac{7}{4}s(b' - kb_n) \\ &\geqslant (s/k)(\gamma_n - \tfrac{7}{4}k^2 b_n) + \tfrac{7}{4}k^2 b' \\ &\geqslant \tfrac{7}{4}k^2(b' - (s/k)^{n+1}) \geqslant \tfrac{7}{8}k^2 b'. \end{aligned} \tag{7.12}$$

But $b' = kb_n + h \leqslant \frac{3}{2}kb_n < \theta^{-1}$, and so it follows from Lemma 6.2 that

$$[[K_{a',b'}(X)]] \ll X^{\eta_{s+k}+\delta}(M^{b'})^{\frac{1}{2}k(k+1)}. \tag{7.13}$$

Thus, on substituting (7.12) and (7.13) into (7.9), we arrive at the upper bound

$$X^{\eta_{s+k}(1+\psi_{n+1}\theta)} \ll X^{\eta_{s+k}+(c_{n+1}+1)\delta}(M^{b'})^{\frac{1}{2}k(k+1)-\frac{7}{8}k^2}.$$

We now recall that $c_{n+1} \leqslant 3(s/k)^{n+1}$ and thus confirm that $X^{(c_{n+1}+1)\delta} < M^{1/2}$. In this way, we obtain the upper bound $X^{\eta_{s+k}\psi_{n+1}\theta} \ll M^{-1/2}$. Since ψ_{n+1} and θ are both positive, we are forced to conclude that $\eta_{s+k} < 0$, contradicting our opening hypothesis. The assumption that $b' \geqslant 2(s/k)^{n+1}$ is therefore untenable, and so we must in fact have $b' < 2(s/k)^{n+1}$. We take h_n to be the integer h at hand, so that $b' = b_{n+1}$ and $\gamma' = \gamma_{n+1}$, and thereby we obtain the

desired conclusion that (7.6) and (7.7) hold with n replaced by $n+1$. This completes the present inductive step.

At this point, we have confirmed the validity of the relations (7.6) and (7.7) for $0 \leqslant n \leqslant N$. We next bound the sequences occurring in (7.7) so as to extract a suitable conclusion. The bound $c_n \leqslant 3(s/k)^n$ has already been confirmed, and the lower bound $\gamma_n \geqslant 0$ already suffices for our purposes at this stage. In addition, the relation (7.2) plainly implies that $b_n \geqslant k^n$, whence from (7.3) we deduce that for $s \geqslant k^2$, one has

$$\psi_{n+1} \geqslant k\psi_n + (k-1)k^n,$$

and by induction this delivers the lower bound $\psi_n \geqslant n(k-1)k^{n-1}$. Finally, we find from (7.6) that $b_N\theta < k/s < 1$, whence $b_N < \theta^{-1}$. Making use of Lemma 6.2, therefore, we find from (7.7) that

$$X^{\eta_{s+k}(1+\psi_N\theta)} \ll X^{\eta_{s+k}+(c_N+1)\delta}(M^{b_N})^{\frac{1}{2}k(k+1)} \ll X^{\eta_{s+k}+k^2}. \tag{7.14}$$

But since $\theta = \frac{1}{2}(k/s)^{N+1}$, it follows that

$$\eta_{s+k} \leqslant \frac{k^2}{\psi_N\theta} \leqslant \frac{2k^2(s/k)^{N+1}}{N(k-1)k^{N-1}}.$$

It is at this point only that we restrict s to be k^2, and thus we obtain the upper bound $\eta_{k(k+1)} \leqslant 2k^4/N$. But we are at liberty to take N as large as we please in terms of k, and thus $\eta_{k(k+1)}$ can be made arbitrarily small. We are therefore forced to conclude that in fact $\eta_{k(k+1)} = 0$. But then, as in the discussion of the opening paragraph of Section 3, we may conclude that $\eta_s = 0$ whenever $s \geqslant k(k+1)$. This completes the proof of the lemma. □

We have now reached the crescendo of this opus, for in view of (2.1) and (2.2), the conclusion of Lemma 7.2 already establishes Theorem 1.1.

A perusal of the proof of Lemma 7.2 might give the impression that it is critical to the success of our iterative process that $s = k^2$, and that the method is inherently unstable. This notion is, however, mistaken. If one were to have $s > \frac{3}{2}k^2$, then one easily reaches the conclusion that $\eta_{s+k} = 0$ simply by comparing the rates of growth of ψ_n and b_n in the above argument. Such a procedure can also be adapted, with care, to the range $s > \frac{5}{4}k^2$. It is only when $k^2 \leqslant s \leqslant \frac{5}{4}k^2$ that the behaviour of the sequences (b_n) and (ψ_n), depending as they do on (h_n), become so difficult to control. The restriction to the case $s = k^2$ should, therefore, be seen rather as a simplifying manoeuvre rather than an inescapable mandate.

8. **Estimates of Weyl type**

The derivation of our upper bounds for Weyl sums, and the application of these estimates to analyse the distribution of polynomials modulo 1, is easily

accomplished by applying Theorem 1.1 within results familiar from the literature. We are therefore concise in our discussion of the associated arguments.

The proof of Theorem 1.5. With the hypotheses of the statement of Theorem 1.5, it follows from [30, Th. 5.2] that for each natural number s, one has

$$f_k(\boldsymbol{\alpha};X) \ll (J_{s,k-1}(2X)X^{\frac{1}{2}k(k-1)}(q^{-1}+X^{-1}+qX^{-j}))^{1/(2s)}\log(2X).$$

But from Theorem 1.1 it follows that when $s=k(k-1)$, one has

$$J_{s,k-1}(2X) \ll X^{2s-\frac{1}{2}k(k-1)+\varepsilon},$$

and thus

$$f_k(\boldsymbol{\alpha};X) \ll X^{1+\varepsilon}(q^{-1}+X^{-1}+qX^{-j})^{1/(2k(k-1))}.$$

As we shall find in Section 9 below, when $s \geqslant k^2-k+1$, one has also the ε-free upper bound $J_{s,k-1}(X) \ll X^{2s-\frac{1}{2}k(k-1)}$, and in like manner this delivers the estimate

$$f_k(\boldsymbol{\alpha};X) \ll X(q^{-1}+X^{-1}+qX^{-j})^{1/(2k^2-2k+2)}\log(2X). \qquad \square$$

The proof of Theorem 1.6. One may establish Theorem 1.6 by applying the argument underlying the proofs of [3, Ths. 4.3 and 4.4]. Let ε be a sufficiently small positive number. We begin by fixing τ to be a positive number with $\tau \leqslant 1/(4k(k-1))-\varepsilon$ and then put $A=X^{1-\tau}$. Observe next that Theorem 1.1 shows that one may replace θ by ε in the case $l=k$ of [3, Th. 4.3]. In this way, we find that the hypotheses of the statement of Theorem 1.6 imply that there exist coprime pairs of integers q_j, b_j $(2 \leqslant j \leqslant k)$ such that

$$q_j \geqslant 1, \quad |q_j\alpha_j - b_j| \leqslant X^{\varepsilon-j}(X/A)^{2k(k-1)} \ (2 \leqslant j \leqslant k)$$

and such that the least common multiple q_0 of $q_2,\dots,q_k$ satisfies

$$q_0 \leqslant X^{\varepsilon}(X/A)^{2k(k-1)}.$$

Notice here that

$$X^{\varepsilon}(X/A)^{2k(k-1)} \leqslant X^{\varepsilon}(X^{\tau})^{2k(k-1)} \leqslant X^{\frac{1}{2}-3\varepsilon}.$$

Write r for q_0 and v_j for b_jq_0/q_j $(2 \leqslant j \leqslant k)$. Then one has

$$|r\alpha_j - v_j| \leqslant X^{2\varepsilon-j}(X/A)^{4k(k-1)} \leqslant X^{1-j}/(4k^4) \quad (2 \leqslant j \leqslant k).$$

Next, denote by d the greatest common divisor $d=(r,v_2,\dots,v_k)$. Then, with the hypotheses of the statement of Theorem 1.6, it is a consequence of [3, Lemma 4.6] that there is a natural number t with $t \leqslant 2k^2$ such that

$$trd^{-1} \leqslant (X/A)^kX^{3k\varepsilon},$$
$$t|r\alpha_j - v_j|d^{-1} \leqslant (X/A)^kX^{3k\varepsilon-j} \ (2 \leqslant j \leqslant k),$$
$$\|trd^{-1}\alpha_1\| \leqslant (X/A)^kX^{3k\varepsilon-1}.$$

But $(X/A)^k = X^{k\tau}$, and so whenever $\delta > k\tau + 3k\varepsilon$, one may conclude that there exist integers $q, a_1, \dots, a_k$ such that

$$1 \leqslant q \leqslant X^{\delta} \quad \text{and} \quad |q\alpha_j - a_j| \leqslant X^{\delta - j} \quad (1 \leqslant j \leqslant k).$$

Since we have supposed ε to be sufficiently small, the same conclusion follows whenever $\delta > k\tau$, and so the proof of Theorem 1.6 is complete. □

The proof of Theorem 1.7. We may apply the argument of the proof of [3, Th. 4.4], substituting the modifications available from Theorem 1.6 above and its proof. Let δ be a positive number. Suppose that $P \ll X$ and $(MXP^{-1})^{4k(k-1)} \leqslant X^{1-\delta}$. Then we find that when

$$\sum_{m=1}^{M} |f_k(m\boldsymbol{\alpha}; X)| \geqslant P,$$

then there exist integers $y, u_1, \dots, u_k$ such that

$$1 \leqslant y \leqslant M(MXP^{-1})^k X^{\varepsilon} \quad \text{and} \quad |y\alpha_j - u_j| \leqslant (MXP^{-1})^k X^{\varepsilon - j} \ (1 \leqslant j \leqslant k).$$

From here, as in the proof of [3, Th. 4.5], the remaining part of our argument is straightforward. If one has

$$\min_{1 \leqslant n \leqslant X} \|\alpha_1 n + \cdots + \alpha_k n^k\| > X^{\delta - \tau(k)}, \tag{8.1}$$

then with $M = [X^{\tau(k) - \delta}] + 1$, one obtains the lower bound

$$\sum_{m=1}^{M} |f_k(m\boldsymbol{\alpha}; X)| > \tfrac{1}{6} X.$$

The above discussion then shows that there exists a natural number y such that

$$y \ll M^{k+1} X^{\varepsilon} \ll X^{(k+1)\tau(k) + \varepsilon} \quad \text{and} \quad \|y\alpha_j\| \ll X^{k\tau(k) - j + \varepsilon} \quad (1 \leqslant j \leqslant k).$$

Thus we find that $y \leqslant X$ and that

$$\|\alpha_1 y + \cdots + \alpha_k y^k\| \leqslant \sum_{j=1}^{k} X^{j-1} \|y\alpha_j\| \ll X^{k\tau(k) - 1 + \varepsilon} < X^{-\tau(k)}.$$

This upper bound contradicts our earlier hypothesis (8.1), and thus we are forced to conclude that

$$\min_{1 \leqslant n \leqslant X} \|\alpha_1 n + \cdots + \alpha_k n^k\| \leqslant X^{\delta - \tau(k)}.$$

This completes the proof of Theorem 1.7. □

9. **Tarry's problem, and related topics**

Our discussion of Tarry's problem follows a familiar path. Let s be a natural number with $s \leqslant k^3$, and define $\rho(\mathbf{h})$ to be the number of integral solutions of the system of equations

$$\sum_{i=1}^{s} x_i^j = h_j \quad (1 \leqslant j \leqslant k),$$

with $1 \leqslant \mathbf{x} \leqslant X$. In addition, let $\sigma(\mathbf{g})$ denote the number of integral solutions of the system of equations

$$\sum_{i=1}^{s} x_i^j = g_j \quad (1 \leqslant j \leqslant k+1),$$

with $1 \leqslant \mathbf{x} \leqslant X$. Observe that

$$\rho(\mathbf{h}) = \sum_{1 \leqslant g_{k+1} \leqslant sX^{k+1}} \sigma(\mathbf{h}, g_{k+1}).$$

Consequently, if for all values of $\mathbf{h}$ one were to have $\sigma(\mathbf{h}, g_{k+1}) \neq 0$ only for a set $\mathcal{A}(\mathbf{h})$ of values of g_{k+1} of cardinality at most t, then it would follow from Cauchy's inequality that

$$\rho(\mathbf{h})^2 \leqslant \Big(\sum_{\substack{1 \leqslant g_{k+1} \leqslant sX^{k+1} \\ g_{k+1} \in \mathcal{A}(\mathbf{h})}} \sigma(\mathbf{h}, g_{k+1}) \Big)^2 \leqslant \operatorname{card}(\mathcal{A}(\mathbf{h})) \sum_{1 \leqslant g_{k+1} \leqslant sX^{k+1}} \sigma(\mathbf{h}, g_{k+1})^2.$$

If such were the case, then one would have

$$\begin{aligned} J_{s,k}(X) &= \sum_{1 \leqslant h_1 \leqslant sX} \cdots \sum_{1 \leqslant h_k \leqslant sX^k} \rho(\mathbf{h})^2 \\ &\leqslant t \sum_{1 \leqslant h_1 \leqslant sX} \cdots \sum_{1 \leqslant h_k \leqslant sX^k} \sum_{1 \leqslant g_{k+1} \leqslant sX^{k+1}} \sigma(\mathbf{h}, g_{k+1})^2 = t J_{s,k+1}(X). \end{aligned}$$

What we have shown is that when X is sufficiently large, and $J_{s,k}(X) > tJ_{s,k+1}(X)$, then there exists a choice of $\mathbf{h}$ such that there are more than t choices for g_{k+1} with $\sigma(\mathbf{h}, g_{k+1}) > 0$. There therefore exists a solution of the system

$$\sum_{i=1}^{s} x_{i1}^j = \sum_{i=1}^{s} x_{i2}^j = \ldots = \sum_{i=1}^{s} x_{it}^j \quad (1 \leqslant j \leqslant k),$$

in which the sums $\sum_{i=1}^{s} x_{il}^{k+1}$ $(1 \leqslant l \leqslant t)$ take distinct values. We have therefore shown that whenever

$$J_{s,k}(X) > tJ_{s,k+1}(X), \tag{9.1}$$

then $W(k,t) \leqslant s$.

We seek to establish that for some positive number δ, one has

$$J_{s,k+1}(X) \ll X^{2s - \frac{1}{2}k(k+1) - \delta}. \tag{9.2}$$

In view of the lower bound (1.5), an estimate of this quality suffices to establish (9.1). But from Theorem 1.1, one has

$$J_{s,k+1}(X) \ll X^{2s-\frac{1}{2}(k+1)(k+2)+\varepsilon}$$

whenever $s \geqslant (k+1)(k+2)$. Moreover, the estimate

$$J_{k+2,k+1}(X) \ll X^{k+2}$$

follows from [31], and indeed earlier results would suffice here. By interpolating via Hölder's inequality, therefore, we find that when s is an integer with $k+2 \leqslant s \leqslant (k+1)(k+2)$, then

$$J_{s,k+1}(X) \ll X^{2s-\frac{1}{2}(k+1)(k+2)+\eta_s+\varepsilon},$$

where

$$\begin{aligned}\eta_s &= ((k+1)(k+2)-s)\left(\frac{\frac{1}{2}(k+1)(k+2)-(k+2)}{(k+1)(k+2)-(k+2)}\right)\\ &= \tfrac{1}{2}(1-1/k)((k+1)(k+2)-s).\end{aligned}$$

It follows that the condition (9.2) is satisfied whenever

$$\tfrac{1}{2}(1-1/k)((k+1)(k+2)-s) < k+1,$$

or equivalently,

$$(k+1)(k+2)-s < 2k\left(\frac{k+1}{k-1}\right) = 2k+4+\frac{4}{k-1}.$$

We deduce that (9.2) holds whenever $s \geqslant (k+1)(k+2)-2k-4$, and hence $W(k,t) \leqslant k^2+k-2$. This completes the proof of Theorem 1.3.

There may be some scope for improvement in the upper bound presented in Theorem 1.3 by exploiting the sharpest bounds available from Vinogradov's mean value theorem for smaller moments (see [37], [39], [10] and [6]). In this way, one might hope to improve even the coefficient of k in the upper bound for $W(k,h)$, though not that of k^2. Of course, in the low degree cases in which $k \leqslant 4$, the above proof of Theorem 1.3 already yields $W(k,t) \leqslant k^2+k-3$. However, stronger conclusions are available in such circumstances.

The proof of Theorem 1.2. Let s and k be natural numbers with $k \geqslant 3$ and $s \geqslant k^2+k+1$, and let X be a positive number sufficiently large in terms of s and k. We follow the argument of the proof of [42, Th. 3]. When $1 \leqslant q \leqslant X^{1/k}$, $1 \leqslant a_j \leqslant q$ $(1 \leqslant j \leqslant k)$ and $(q,a_1,\dots,a_k)=1$, define the major arc $\mathfrak{M}(q,\mathbf{a})$ by

$$\mathfrak{M}(q,\mathbf{a}) = \{\boldsymbol{\alpha} \in [0,1)^k : |q\alpha_j - a_j| \leqslant X^{1/k-j}\ (1 \leqslant j \leqslant k)\}.$$

It is not hard to check that the arcs $\mathfrak{M}(q,\mathbf{a})$ are disjoint. Let $\mathfrak{M}$ denote the union of the major arcs $\mathfrak{M}(q,\mathbf{a})$ with q and $\mathbf{a}$ as above, and define the minor arcs $\mathfrak{m}$ by $\mathfrak{m}=[0,1)^k\setminus\mathfrak{M}$. Then from (1.2), we have

$$J_{s,k}(X)=\int_{\mathfrak{M}}|f(\boldsymbol{\alpha};X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha}+\int_{\mathfrak{m}}|f(\boldsymbol{\alpha};X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha}. \tag{9.3}$$

We first bound the contribution of the minor arcs. As a consequence of Theorem 1.6, one finds that

$$\sup_{\boldsymbol{\alpha}\in\mathfrak{m}}|f(\boldsymbol{\alpha};X)|\leqslant X^{1-\tau+\varepsilon},$$

where $\tau^{-1}=4k(k-1)$. Then it follows from Theorem 1.1 that

$$\begin{aligned}\int_{\mathfrak{m}}|f(\boldsymbol{\alpha};X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha}&\ll\Big(\sup_{\boldsymbol{\alpha}\in\mathfrak{m}}|f(\boldsymbol{\alpha};X)|\Big)^{2s-2k^2-2k}\oint|f(\boldsymbol{\alpha};X)|^{2k^2+2k}\,\mathrm{d}\boldsymbol{\alpha}\\&\ll(X^{1-\tau+\varepsilon})^{2s-2k^2-2k}X^{\frac{3}{2}k(k+1)+\varepsilon}\\&\ll X^{2s-\frac{1}{2}k(k+1)-1/(3k^2)}.\end{aligned} \tag{9.4}$$

Next we discuss the major arc contribution. When $\boldsymbol{\alpha}\in\mathfrak{M}(q,\mathbf{a})\subseteq\mathfrak{M}$, write

$$V(\boldsymbol{\alpha};q,\mathbf{a})=q^{-1}S(q,\mathbf{a})I(\boldsymbol{\alpha}-\mathbf{a}/q;X),$$

where

$$S(q,\mathbf{a})=\sum_{r=1}^{q}e((a_1r+\cdots+a_kr^k)/q)$$

and

$$I(\boldsymbol{\beta};X)=\int_0^X e(\beta_1\gamma+\cdots+\beta_k\gamma^k)\,\mathrm{d}\gamma.$$

In addition, define the function $V(\boldsymbol{\alpha})$ to be $V(\boldsymbol{\alpha};q,\mathbf{a})$ when $\boldsymbol{\alpha}\in\mathfrak{M}(q,\mathbf{a})\subseteq\mathfrak{M}$ and to be zero otherwise. Then the argument concluding [42, §3] shows that

$$\begin{aligned}&\int_{\mathfrak{M}}|f(\boldsymbol{\alpha};X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha}-\int_{\mathfrak{M}}|V(\boldsymbol{\alpha})|^{2s}\,\mathrm{d}\boldsymbol{\alpha}\\&\qquad\ll X^{1+2/k}\Big(\oint|f(\boldsymbol{\alpha};X)|^{2s-2}\,\mathrm{d}\boldsymbol{\alpha}+\oint|V(\boldsymbol{\alpha})|^{2s-2}\,\mathrm{d}\boldsymbol{\alpha}\Big).\end{aligned} \tag{9.5}$$

When $\boldsymbol{\alpha}\in\mathfrak{M}(q,\mathbf{a})\subseteq\mathfrak{M}$, one has $(q,a_1,\dots,a_k)=1$ and $|q\alpha_j-a_j|\leqslant X^{1/k-j}$ $(1\leqslant j\leqslant k)$. Then it follows from [30, Ths. 7.1 and 7.3] that when $\boldsymbol{\alpha}\in\mathfrak{M}(q,\mathbf{a})\subseteq\mathfrak{M}$, one has

$$V(\boldsymbol{\alpha})\ll Xq^{\varepsilon}(q+|q\alpha_1-a_1|X+\cdots+|q\alpha_k-a_k|X^k)^{-1/k}.$$

Consequently, one finds that when $t\geqslant\frac{1}{2}k(k+1)$, one has

$$\int_{\mathfrak{M}}|V(\boldsymbol{\alpha})|^{2t}\,\mathrm{d}\boldsymbol{\alpha}\ll X^{2t}WZ,$$

where

$$W = \sum_{1\leqslant q\leqslant X^{1/k}} \sum_{a_1=1}^{q} \cdots \sum_{a_k=1}^{q} (q^{\varepsilon-1/k})^{2t}$$

and

$$Z = \prod_{j=1}^{k} \int_0^{X^{1/k-j}} (1+\beta_j X^j)^{-2t/k^2}\,\mathrm{d}\beta_j.$$

But since $2t \geqslant k(k+1)$, we obtain the upper bounds

$$W \ll X^{1/(3k)} \sum_{q=1}^{\infty} q^{-5/4} \ll X^{1/(3k)} \tag{9.6}$$

and

$$Z \ll \prod_{j=1}^{k} \int_0^{\infty} (1+\beta_j X^j)^{-1-1/k}\,\mathrm{d}\beta_j \ll X^{-\frac{1}{2}k(k+1)}. \tag{9.7}$$

Thus, in particular, we deduce that when $s \geqslant k^2+k+1$, then

$$\int_{\mathfrak{M}} |V(\boldsymbol{\alpha})|^{2s-2}\,\mathrm{d}\boldsymbol{\alpha} \ll X^{2s-2-\frac{1}{2}k(k+1)+1/(3k)}.$$

In combination with Theorem 1.1, this leads from (9.5) to the asymptotic relation

$$\int_{\mathfrak{M}} |f(\boldsymbol{\alpha};X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha} - \int_{\mathfrak{M}} |V(\boldsymbol{\alpha})|^{2s}\,\mathrm{d}\boldsymbol{\alpha} \ll X^{2s-\frac{1}{2}k(k+1)-1/(3k)}. \tag{9.8}$$

The argument employed in deriving (9.6) and (9.7) is readily adapted to show that the singular series $\mathfrak{S}(s,k)$ defined in (1.8), and the singular integral $\mathfrak{J}(s,k)$ defined in (1.9), both converge absolutely, and that

$$\int_{\mathfrak{M}} |V(\boldsymbol{\alpha})|^{2s}\,\mathrm{d}\boldsymbol{\alpha} = \mathfrak{S}(s,k)\mathfrak{J}(s,k) + O(X^{2s-\frac{1}{2}k(k+1)-1/(3k)}).$$

The asymptotic formula claimed implicitly in Theorem 1.2 now follows by substituting (9.4) and (9.8) into (9.3). This completes the proof of Theorem 1.2. □

As essentially was observed by Vaughan, one must have both $\mathfrak{S}(s,k) \gg 1$ and $\mathfrak{J}(s,k) \gg 1$ (see the conclusion of [30, §7.3]). For otherwise one would have

$$\oint |f(\boldsymbol{\alpha};X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha} = o(X^{2s-\frac{1}{2}k(k+1)}),$$

which contradicts the elementary lower bound (1.5).

An argument similar to that employed in the proof of Theorem 1.2 delivers an asymptotic formula for the number of solutions of a more general diagonal

Diophantine system. When s and k are natural numbers, and a_{ij} are integers for $1 \leqslant i \leqslant k$ and $1 \leqslant j \leqslant s$, we write

$$\phi_i(\mathbf{x}) = \sum_{j=1}^{s} a_{ij} x_j^i \quad (1 \leqslant i \leqslant k),$$

and we consider the Diophantine system

$$\phi_i(\mathbf{x}) = 0 \quad (1 \leqslant i \leqslant k). \tag{9.9}$$

We write $N(B)$ for the number of integral solutions of the system (9.9) with $|\mathbf{x}| \leqslant B$. We next define the (formal) real and p-adic densities associated with the system (9.9), and here we follow Schmidt [25]. When $L > 0$, define

$$\lambda_L(\eta) = \begin{cases} L(1 - L|\eta|), & \text{when } |\eta| \leqslant L^{-1}, \\ 0, & \text{otherwise.} \end{cases}$$

We then put

$$\mu_L = \int_{|\boldsymbol{\xi}| \leqslant 1} \prod_{i=1}^{k} \lambda_L(\phi_i(\boldsymbol{\xi}))\, \mathrm{d}\boldsymbol{\xi}.$$

The limit $\sigma_\infty = \lim_{L \to \infty} \mu_L$, when it exists, is called the *real density*. Meanwhile, given a natural number q, we write

$$M(q) = \operatorname{card}\{\mathbf{x} \in (\mathbb{Z}/q\mathbb{Z})^s \,:\, \phi_i(\mathbf{x}) \equiv 0 \pmod{q}\ (1 \leqslant i \leqslant k)\}.$$

For each prime number p, we then put

$$\sigma_p = \lim_{H \to \infty} p^{H(k-s)} M(p^H),$$

provided that this limit exists, and refer to σ_p as the *p-adic density*.

THEOREM 9.1. *Let s and k be natural numbers with $k \geqslant 3$ and $s \geqslant 2k^2 + 2k + 1$. Suppose that a_{ij} $(1 \leqslant i \leqslant k,\ 1 \leqslant j \leqslant s)$ are nonzero integers. Suppose, in addition, that the system of equations* (9.9) *possess nonsingular real and p-adic solutions for each prime number p. Then one has*

$$N(B) \sim \sigma_\infty \Big(\prod_p \sigma_p \Big) B^{s - \frac{1}{2}k(k+1)}.$$

In particular, the system (9.9) *satisfies the Hasse Principle.*

We will not offer any details of the proof here, the argument following in most respects that of the proof of Theorem 1.2. We note only that the system (9.9), if singular, is easily shown to have a singular locus of affine dimension at most $k-1$, which is harmless in the analysis. We note also that the restriction that $a_{ij} \neq 0$ $(1 \leqslant i \leqslant k,\ 1 \leqslant j \leqslant s)$ may be largely removed by elaborating on the basic argument. We emphasise that the most striking feature of Theorem 9.1 is that such a conclusion cannot possibly hold when $s < \frac{1}{2}k(k+1)$. Thus, for the very first time for a system of diagonal equations of

higher degree, we have an asymptotic formula in which the number of variables is just four times the best possible result. Hitherto, the number of variables required to achieve a successful analysis would be roughly $2\log k$ times the best possible result, a factor which becomes arbitrarily large as k increases.

We turn our attention next to the Hilbert-Kamke problem, a generalisation of Waring's problem considered first by Hilbert [13]. When $n_1,\dots,n_k$ are natural numbers, let $R_{s,k}(\mathbf{n})$ denote the number of solutions in natural numbers $\mathbf{x}$ of the system of equations

$$\sum_{i=1}^{s} x_i^j = n_j \quad (1 \leqslant j \leqslant k). \tag{9.10}$$

Put

$$X = \max_{1\leqslant j\leqslant k} n_j^{1/j},$$

and then write

$$\mathcal{J}_{s,k}(\mathbf{n}) = \int_{\mathbb{R}^k} I(\boldsymbol{\beta};1)^s e(-\beta_1 n_1/X - \cdots - \beta_k n_k/X^k)\,\mathrm{d}\boldsymbol{\beta}$$

and

$$\mathcal{S}_{s,k}(\mathbf{n}) = \sum_{q=1}^{\infty} \sum_{\substack{1\leqslant \mathbf{a}\leqslant q\\ (q,a_1,\dots,a_k)=1}} (q^{-1}S(q,\mathbf{a}))^s e(-(a_1n_1+\cdots+a_kn_k)/q).$$

The local solubility conditions associated with the system (9.10) are quite subtle, and we refer the reader to [1] for a discussion of the conditions under which real and p-adic solutions may be expected to exist for the system (9.10). It is easy to see, however, that the conditions

$$n_k^{j/k} \leqslant n_j \leqslant s^{1-j/k} n_k^{j/k} \quad (1 \leqslant j \leqslant k)$$

are needed. One also finds that p-adic solubility is not assured without at least 2^k variables.

THEOREM 9.2. *Let s and k be natural numbers with $k \geqslant 3$ and $s \geqslant 2k^2+2k+1$. Suppose that the natural numbers $n_1,\dots,n_k$ are sufficiently large in terms of s and k. Put $X = \max_{1\leqslant j\leqslant k} n_j^{1/j}$. Suppose, in addition, that the system* (9.10) *has nonsingular real and p-adic solutions. Then one has*

$$R_{s,k}(\mathbf{n}) = \mathcal{J}_{s,k}(\mathbf{n})\mathcal{S}_{s,k}(\mathbf{n})X^{s-\frac{1}{2}k(k+1)} + o(X^{s-\frac{1}{2}k(k+1)}).$$

We refer the reader to [1], [20], [21] for the many details associated with a successful treatment of this problem. The technology available at the time of writing of the latter papers made necessary the constraint $s \geqslant (4+o(1))k^2\log k$ in place of the lower bound $s \geqslant 2k^2+2k+1$ in Theorem 9.2. Our observation here is that a successful local-global analysis is now available via the circle

method when the number of variables grows like $2k^2+2k+1$, only a factor of 4 away from what is likely to be best possible.

10. **The asymptotic formula in Waring's problem**

The proof of Theorem 1.4 would be routine were our goal the less precise bound $\widetilde{G}(k) \leqslant 2k^2+2k+1$. Saving four additional variables requires some discussion that hints at possible new strategies for transforming estimates for $J_{s,k}(X)$ into upper bounds for $\widetilde{G}(k)$. En route, we also improve some old estimates of Hua [17].

Write

$$g(\alpha) = \sum_{1\leqslant x\leqslant X} e(\alpha x^k),$$

and when $s \in \mathbb{N}$, define

$$I_s(X) = \int_0^1 |g(\alpha)|^{2s}\,\mathrm{d}\alpha.$$

Then on considering the underlying Diophantine system, one has

$$\begin{aligned} I_s(X) &= \sum_{|h_1|\leqslant sX} \cdots \sum_{|h_{k-1}|\leqslant sX^{k-1}} \oint |f(\boldsymbol{\alpha};X)|^{2s} e(-h_1\alpha_1-\cdots-h_{k-1}\alpha_{k-1})\,\mathrm{d}\boldsymbol{\alpha} \\ &\ll X^{\frac{1}{2}k(k-1)} \oint |f(\boldsymbol{\alpha};X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha} = X^{\frac{1}{2}k(k-1)} J_{s,k}(X). \end{aligned}$$

Thus we obtain the classical bound

$$I_s(X) \ll X^{2s-k+\eta_s+\varepsilon}. \tag{10.1}$$

Ford [9] obtained a bound potentially sharper, valid for each natural number m with $1 \leqslant m \leqslant k$, and $s \geqslant \frac{1}{2}m(m-1)$, which is tantamount to

$$I_s(X) \ll X^{2s-k+\eta^*_{s,m}+\varepsilon},$$

where $\eta^*_{s,m} = \frac{1}{m}\eta_{s-\frac{1}{2}m(m-1)}$. A little later, this conclusion was obtained independently by Ustinov [27]. Owing to the efficiency of Theorem 1.1, this estimate proves to be no sharper than that provided by (10.1), at least in applications to the asymptotic formula in Waring's problem. Instead we offer a very modest refinement of (10.1). The idea underlying this refinement is related to one first shown to the author by Bob Vaughan in the first year of the author's Ph.D. studies, in 1988.

Lemma 10.1. *For each natural number s, one has*

$$I_s(X) \ll X^{\varepsilon}(X^{2s-k-1+\eta_{s,k}} + X^{2s-k+\eta_{s,k-1}}).$$

Proof. Define the exponential sum $F(\boldsymbol{\beta}) = F_k(\boldsymbol{\beta}; X)$ by

$$F(\boldsymbol{\beta}) = \sum_{1 \leqslant x \leqslant X} e(\beta_k x^k + \beta_{k-2} x^{k-2} + \cdots + \beta_1 x).$$

Thus, to be precise, the argument of the exponentials in $F(\boldsymbol{\beta})$ is a polynomial of degree k in which the coefficient of the monomial of degree $k-1$ is zero. Also, define $\Upsilon_k(X; h)$ to be the number of integral solutions of the Diophantine system

$$\begin{aligned} \sum_{i=1}^{s} (x_i^j - y_i^j) &= 0 \quad (1 \leqslant j \leqslant k,\ j \neq k-1), \\ \sum_{i=1}^{s} (x_i^{k-1} - y_i^{k-1}) &= h, \end{aligned} \tag{10.2}$$

with $1 \leqslant \mathbf{x}, \mathbf{y} \leqslant X$. Then on considering the underlying Diophantine system, one finds that

$$\oint |F(\boldsymbol{\beta})|^{2s} \, \mathrm{d}\boldsymbol{\beta} = \sum_{|h| \leqslant sX^{k-1}} \Upsilon_k(X; h). \tag{10.3}$$

By applying an integer shift z to the variables in the system (10.2), we find that $\Upsilon_k(X; h)$ counts the number of integral solutions of the Diophantine system

$$\begin{aligned} \sum_{i=1}^{s} ((x_i - z)^j - (y_i - z)^j) &= 0 \quad (1 \leqslant j \leqslant k,\ j \neq k-1), \\ \sum_{i=1}^{s} ((x_i - z)^{k-1} - (y_i - z)^{k-1}) &= h, \end{aligned}$$

with $1 + z \leqslant \mathbf{x}, \mathbf{y} \leqslant X + z$. But by applying the Binomial Theorem, we find that $\mathbf{x}, \mathbf{y}$ satisfies this system of equations if and only if

$$\begin{aligned} \sum_{i=1}^{s} (x_i^j - y_i^j) &= 0 \quad (1 \leqslant j \leqslant k-2) \\ \sum_{i=1}^{s} (x_i^{k-1} - y_i^{k-1}) &= h, \\ \sum_{i=1}^{s} (x_i^k - y_i^k) &= khz. \end{aligned} \tag{10.4}$$

If we restrict the shifts z to lie in the interval $1 \leqslant z \leqslant X$, then we see that an upper bound for $\Upsilon_k(X; h)$ is given by the number of integral solutions of the system (10.4) with $1 \leqslant \mathbf{x}, \mathbf{y} \leqslant 2X$. On considering the underlying Diophantine system, we therefore deduce from (10.3) that for each integer z with $1 \leqslant z \leqslant X$,

one has

$$\oint |F(\boldsymbol{\beta})|^{2s}\,\mathrm{d}\boldsymbol{\beta} \leqslant \sum_{|h|\leqslant sX^{k-1}} \oint |f(\boldsymbol{\alpha};2X)|^{2s} e(-(kz\alpha_k+\alpha_{k-1})h)\,\mathrm{d}\boldsymbol{\alpha}.$$

Hence

(10.5)
$$\begin{aligned}\oint |F(\boldsymbol{\beta})|^{2s}\,\mathrm{d}\boldsymbol{\beta} &\ll X^{-1} \sum_{1\leqslant z\leqslant X} \oint |f(\boldsymbol{\alpha};2X)|^{2s} \min\{X^{k-1}, \|kz\alpha_k+\alpha_{k-1}\|^{-1}\}\,\mathrm{d}\boldsymbol{\alpha}\\ &= X^{-1}\oint |f(\boldsymbol{\alpha};2X)|^{2s}\Psi(\alpha_k,\alpha_{k-1})\,\mathrm{d}\boldsymbol{\alpha},\end{aligned}$$

where we have written

$$\Psi(\alpha_k,\alpha_{k-1}) = \sum_{1\leqslant z\leqslant X} \min\{X^{k-1}, \|kz\alpha_k+\alpha_{k-1}\|^{-1}\}.$$

Suppose that $\alpha_k \in \mathbb{R}$ and that $b\in\mathbb{Z}$ and $r\in\mathbb{N}$ satisfy $(b,r)=1$ and $|\alpha_k - b/r| \leqslant r^{-2}$. Then it follows from [3, Lemma 3.2][4] that

$$\begin{aligned}\Psi(\alpha_k,\alpha_{k-1}) &\ll (X^{k-1}+r\log(2r))(X/r+1)\\ &\ll X^k(X^{-1}+r^{-1}+rX^{-k})(\log(2r)).\end{aligned} \tag{10.6}$$

Applying a standard transference principle (compare Exercise 2 of [30, §2.8]), it follows that

$$\Psi(\alpha_k,\alpha_{k-1}) \ll X^{k+\varepsilon}(X^{-1}+(r+X^k|r\alpha_k-b|)^{-1}+(r+X^k|r\alpha_k-b|)X^{-k}). \tag{10.7}$$

We now return to consider the relation (10.5). Let $\mathfrak{m}$ denote the set of real numbers $\alpha\in[0,1)$ having the property that whenever $q\in\mathbb{N}$ and $\|q\alpha\|\leqslant X^{1-k}$, then $q>X$. Also, let $\mathfrak{M}$ denote the complementary set $[0,1)\setminus\mathfrak{m}$. By Dirichlet's theorem on Diophantine approximation, whenever $\alpha_k\in\mathfrak{m}$, there exists $q\in\mathbb{N}$ with $q\leqslant X^{k-1}$ such that $\|q\alpha_k\|\leqslant X^{1-k}$. From the definition of $\mathfrak{m}$, one must have $q>X$, and hence it follows from (10.6) that

$$\sup_{\alpha_k\in\mathfrak{m}} \Psi(\alpha_k,\alpha_{k-1}) \ll X^{k-1+\varepsilon}.$$

Thus we deduce from (1.2) that

$$\begin{aligned}\int_{\mathfrak{m}\times[0,1)^{k-1}} |f(\boldsymbol{\alpha};2X)|^{2s}\Psi(\alpha_k,\alpha_{k-1})\,\mathrm{d}\boldsymbol{\alpha} &\ll X^{k-1+\varepsilon}\oint |f(\boldsymbol{\alpha};2X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha}\\ &\ll X^{k-1+\varepsilon}J_{s,k}(2X).\end{aligned}$$

[4]We note that the strict inequality $|\alpha_k - b/r| < r^{-2}$ imposed by Baker is unnecessary in the proof of [3, Lemma 3.2]

Substituting this conclusion into (10.5), we see that

$$\oint |F(\boldsymbol{\beta})|^{2s}\,\mathrm{d}\boldsymbol{\beta} \ll X^{k-2+\varepsilon}J_{s,k}(2X) + X^{-1}\int_{\mathfrak{M}\times[0,1)^{k-1}} |f(\boldsymbol{\alpha};2X)|^{2s}\Psi(\alpha_k,\alpha_{k-1})\,\mathrm{d}\boldsymbol{\alpha}. \tag{10.8}$$

Let $\mathfrak{M}(q,a)$ denote the set of real numbers $\alpha_k \in [0,1)$ with $|q\alpha_k - a| \leqslant X^{1-k}$. Then $\mathfrak{M}$ is the union of the sets $\mathfrak{M}(q,a)$ with $0 \leqslant a \leqslant q \leqslant X$ and $(a,q)=1$. From (10.7) it follows that when $\alpha_k \in \mathfrak{M}(q,a) \subseteq \mathfrak{M}$, one has

$$\Psi(\alpha_k,\alpha_{k-1}) \ll X^{k-1+\varepsilon} + X^{k+\varepsilon}(q+X^k|q\alpha_k - a|)^{-1}.$$

Define the function $\Phi(\theta)$ for $\theta \in \mathfrak{M}$ by putting

$$\Phi(\theta) = (q+X^k|q\theta - a|)^{-1}$$

when $\theta \in \mathfrak{M}(q,a) \subseteq \mathfrak{M}$. Then we deduce from (10.8) that

$$\oint |F(\boldsymbol{\beta})|^{2s}\,\mathrm{d}\boldsymbol{\beta} \ll X^{k-2+\varepsilon}J_{s,k}(2X) + X^{k-1+\varepsilon}\mathcal{T}, \tag{10.9}$$

where

$$\mathcal{T} = \int_{\mathfrak{M}} \Phi(\alpha_k) \oint |f(\boldsymbol{\beta},\alpha_k;2X)|^{2s}\,\mathrm{d}\boldsymbol{\beta}\,\mathrm{d}\alpha_k.$$

From Brüdern [7, Lemma 2], we find that

$$\int_{\mathfrak{M}} \Phi(\alpha_k)|f(\boldsymbol{\beta},\alpha_k;2X)|^{2s}\,\mathrm{d}\alpha_k \ll X^{\varepsilon-k}\Big(X\int_0^1 |f(\boldsymbol{\beta},\alpha_k;2X)|^{2s}\,\mathrm{d}\alpha_k + |f(\boldsymbol{\beta},0;2X)|^{2s}\Big),$$

and hence

$$\begin{aligned}\mathcal{T} &\ll X^{\varepsilon-k}\Big(X\oint |f_k(\boldsymbol{\alpha};2X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha} + \oint |f_{k-1}(\boldsymbol{\beta};2X)|^{2s}\,\mathrm{d}\boldsymbol{\beta}\Big)\\ &\ll X^{\varepsilon-k}(XJ_{s,k}(2X) + J_{s,k-1}(2X)).\end{aligned}$$

Consequently, from (10.9) we conclude that

$$\oint |F(\boldsymbol{\beta})|^{2s}\,\mathrm{d}\boldsymbol{\beta} \ll X^{k-2+\varepsilon}J_{s,k}(2X) + X^{\varepsilon-1}J_{s,k-1}(2X). \tag{10.10}$$

Next we observe that, on considering the underlying Diophantine system, one has

$$I_s(X) = \sum_{|h_1|\leqslant sX} \cdots \sum_{|h_{k-2}|\leqslant sX^{k-2}} R(X;\mathbf{h}),$$

where $R(X;\mathbf{h})$ denotes the number of integral solutions of the system

$$\sum_{i=1}^{s}(x_i^j-y_i^j)=h_j \quad (1\leqslant j\leqslant k-2),$$
$$\sum_{i=1}^{s}(x_i^k-y_i^k)=0,$$

with $1\leqslant \mathbf{x},\mathbf{y}\leqslant X$. Thus, again considering the underlying Diophantine system, we obtain the upper bound

$$\begin{aligned} I_s(X) &\ll \sum_{|h_1|\leqslant sX}\cdots\sum_{|h_{k-2}|\leqslant sX^{k-2}}\oint |F(\boldsymbol{\beta})|^{2s}e(-\beta_1h_1-\cdots-\beta_{k-2}h_{k-2})\,\mathrm{d}\boldsymbol{\beta}\\ &\ll X^{\frac{1}{2}(k-1)(k-2)}\oint |F(\boldsymbol{\beta})|^{2s}\,\mathrm{d}\boldsymbol{\beta}. \end{aligned}$$

In view of (10.10), we therefore arrive at the estimate

$$\begin{aligned} I_s(X) &\ll X^{\frac{1}{2}(k+1)(k-2)+\varepsilon}J_{s,k}(2X)+X^{\frac{1}{2}(k-1)(k-2)-1+\varepsilon}J_{s,k-1}(2X)\\ &\ll X^{2s-k-1+\eta_{s,k}+\varepsilon}+X^{2s-k+\eta_{s,k-1}+\varepsilon}. \end{aligned}$$

This completes the proof of the lemma. □

From Theorem 1.1, we have $\eta_{s,k-1}=0$ for $s\geqslant k(k-1)$. By Hölder's inequality, moreover, one finds from Theorem 1.1 that

$$\begin{aligned} \oint |f_k(\boldsymbol{\alpha};X)|^{2k^2+2k-4}\,\mathrm{d}\boldsymbol{\alpha} &\leqslant \Big(\oint |f_k(\boldsymbol{\alpha};X)|^{2k^2+2k}\,\mathrm{d}\boldsymbol{\alpha}\Big)^{1-2/(k^2+k)}\\ &\ll \Big(X^{\frac{3}{2}(k^2+k)+\varepsilon}\Big)^{1-2/(k^2+k)}\\ &\ll X^{\frac{3}{2}(k^2+k)-3+\varepsilon}. \end{aligned}$$

Consequently, one has $\eta_{s,k}\leqslant 1$ for $s\geqslant k^2+k-2$. Then by Lemma 10.1, we obtain the following corollary to Theorem 1.1.

COROLLARY 10.2. *When $s\geqslant k^2+k-2$, one has $I_s(X)\ll X^{2s-k+\varepsilon}$.*

Having prepared the ground, the proof of Theorem 1.4 is now swift. Consider a large integer n, put $X=[n^{1/k}]$ and recall the definition of the sets of arcs $\mathfrak{m}$ and $\mathfrak{M}$ from the proof of Lemma 10.1. From Corollary 10.2 and Weyl's inequality (see [30, Lemma 2.4]), one finds that when $t\geqslant 2k^2+2k-3$, one has

$$\begin{aligned} \int_{\mathfrak{m}} g(\alpha)^t e(-n\alpha)\,\mathrm{d}\alpha &\ll \Big(\sup_{\alpha\in\mathfrak{m}}|g(\alpha)|\Big)^{t-(2k^2+2k-4)}\int_0^1 |g(\alpha)|^{2k^2+2k-4}\,\mathrm{d}\alpha\\ &\ll (X^{1-2^{1-k}+\varepsilon})^{t-(2k^2+2k-4)}X^{(2k^2+2k-4)-k}\\ &\ll X^{t-k-2^{-k}}. \end{aligned}$$

Notice here, of course, that we could have employed the conclusion of Theorem 1.5 in place of Weyl's inequality. Meanwhile, the methods of [30, §4.4] show that, under the same conditions on t, one has

$$\int_{\mathfrak{M}} g(\alpha)^t e(-n\alpha)\,d\alpha \sim \frac{\Gamma(1+1/k)^t}{\Gamma(t/k)}\mathfrak{S}_{t,k}(n)n^{t/k-1} + o(n^{t/k-1}),$$

where $\mathfrak{S}_{t,k}(n)$ is defined as in (1.12). Thus we deduce that for $t \geqslant 2k^2+2k-3$, one has

$$\begin{aligned} R_{t,k}(n) &= \int_{\mathfrak{M}} g(\alpha)^t e(-n\alpha)\,\mathrm{d}\alpha + \int_{\mathfrak{m}} g(\alpha)^t e(-n\alpha)\,\mathrm{d}\alpha \\ &= \frac{\Gamma(1+1/k)^t}{\Gamma(t/k)}\mathfrak{S}_{t,k}(n)n^{t/k-1} + o(n^{t/k-1}), \end{aligned}$$

whence $\widetilde{G}(k) \leqslant 2k^2+2k-3$. This completes the proof of Theorem 1.4.

We take this opportunity to point out that L.-K. Hua investigated the problem of bounding the least integer C_k such that, whenever $s \geqslant C_k$, one has

$$\oint |f_k(\boldsymbol{\alpha}; X)|^s\,\mathrm{d}\boldsymbol{\alpha} \ll X^{s-\frac{1}{2}k(k+1)+\varepsilon},$$

and likewise the least integer S_k such that, whenever $s \geqslant S_k$, one has

$$\oint |F_k(\boldsymbol{\beta}; X)|^s\,\mathrm{d}\boldsymbol{\beta} \ll X^{s-\frac{1}{2}(k^2-k+2)+\varepsilon},$$

pursuing in particular the situation for smaller values of k. His arguments involve a clever application of Weyl differencing in a style that we would describe in the single equation situation as underlying Hua's lemma. In Chapter 5 of [17], one finds tables recording the upper bounds

$$C_3 \leqslant 16, \quad C_4 \leqslant 46, \quad C_5 \leqslant 110, \dots$$

and

$$S_3 \leqslant 10, \quad S_4 \leqslant 32, \quad S_5 \leqslant 86, \dots .$$

The conclusion of Theorem 1.1 shows that $C_k \leqslant 2k(k+1)$, an upper bound superior to the conclusions of Hua for $k \geqslant 4$. Meanwhile, as a consequence of the estimate (10.10), one obtains the estimate contained in the following theorem.

THEOREM 10.3. *Suppose that $k \geqslant 3$ and $s \geqslant k^2+k-2$. Then one has*

$$\oint |F_k(\boldsymbol{\beta}; X)|^{2s}\,\mathrm{d}\boldsymbol{\beta} \ll X^{2s-\frac{1}{2}(k^2-k+2)+\varepsilon}.$$

Proof. The discussion leading to Corollary 10.2 shows that $\eta_{s,k-1} = 0$ for $s \geqslant k(k-1)$ and $\eta_{s,k} \leqslant 1$ for $s \geqslant k^2+k-2$. The desired conclusion is therefore immediate from (3.8), (3.11) and (10.10). □

Thus we have $S_k \leqslant 2k^2+2k-4$, an upper bound superior to those of Hua for $k \geqslant 5$.

11. **A heuristic argument**

We take the opportunity in this section to discuss a heuristic argument that delivers the bound

$$J_{s+k}(X) \ll X^{2s+2k-\frac{1}{2}k(k+1)+\varepsilon} \tag{11.1}$$

for $s \geqslant \frac{1}{2}k(k+1)$. In view of the lower bound (1.5), of course, the bound (11.1) cannot hold for $s+k < \frac{1}{2}k(k+1)$, so is in a strong sense close to best possible.

Our starting point is a heuristic interpretation of Lemma 6.1. In the course of the proof of Lemma 6.1, a critical role is played by the interpretation of the system of equations (6.1) by means of the implied congruences (6.3). In some sense, for each fixed choice of $\mathbf{y}$ in (6.3), the conclusion of Lemma 4.1 indicates that there are at most $k!p^{\frac{1}{2}k(k-1)(a+b)}$ possible choices for $\mathbf{x}$ with $1 \leqslant \mathbf{x} \leqslant p^{kb}$ and $\mathbf{x} \equiv \boldsymbol{\xi} \pmod{p^{a+1}}$ for some $\boldsymbol{\xi} \in \Xi_a(\xi)$. This is transformed via Cauchy's inequality into the statement that, with a compensating factor $k!p^{\frac{1}{2}k(k-1)(a+b)}$, the variables in (6.1) are constrained by the additional congruence relations $\mathbf{x} \equiv \mathbf{y} \pmod{p^{kb}}$. Such an interpretation is embodied in the relation (6.6).

An alternative interpretation, which we emphasise is heuristic in nature and not a statement of fact, is that, by relabelling variables if necessary, the congruences (6.3) essentially amount in (6.1) to the constraint $x_j \equiv y_j \pmod{p^{jb}}$ $(1 \leqslant j \leqslant k)$, with an additional compensating factor of $k!p^{\frac{1}{2}k(k-1)a}$. Indeed, one can prove the initial statement that $x_j \equiv y_j \pmod{p^b}$ $(1 \leqslant j \leqslant k)$ with precisely this compensating factor. Then, by fixing the variables x_1, y_1, and considering the system (6.3) with $2 \leqslant j \leqslant k$, one might suppose that a corresponding constraint $x_j \equiv y_j \pmod{p^{2b}}$ $(2 \leqslant j \leqslant k)$ might be imposed. Then, by fixing the variables x_2, y_2, and considering the system (6.3) with $3 \leqslant j \leqslant k$, one seeks a corresponding constraint $x_j \equiv y_j \pmod{p^{3b}}$ $(3 \leqslant j \leqslant k)$, and so on. Such a heuristic implies a new relation to replace (6.6) of the shape

$$K_{a,b}^{\boldsymbol{\sigma};\boldsymbol{\tau}}(X;\xi,\eta) \ll M^{\frac{1}{2}k(k-1)a} \sum_{\substack{1\leqslant\zeta_1\leqslant p^b\\ \zeta_1\equiv\xi \pmod{p^a}}} \cdots \sum_{\substack{1\leqslant\zeta_k\leqslant p^{kb}\\ \zeta_k\equiv\xi \pmod{p^a}}} \mathcal{I}(\boldsymbol{\zeta}),$$

where

$$\mathcal{I}(\boldsymbol{\zeta}) = \oint \Big(\prod_{i=1}^{k} |\mathfrak{f}_{ib}(\boldsymbol{\alpha};\zeta_i)|^2 \Big) |\mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)|^{2u} \, \mathrm{d}\boldsymbol{\alpha}.$$

Such an assertion at least carries the weight of correctly accounting for the number of available residue classes, though of course one cannot hope for the implied degree of independence to be true in anything but an average sense.

From here, an application of Hölder's inequality leads to the bound

$$K_{a,b}^{\boldsymbol{\sigma};\boldsymbol{\tau}}(X;\xi,\eta) \ll M^{\frac{1}{2}k(k-1)a} \Big(\prod_{i=1}^{k} M^{ib-a} \Theta_{ib,b}(X;\eta)^{1/k} \Big), \tag{11.2}$$

where

$$\Theta_{c,b}(X;\eta) = \max_{1\leqslant\zeta\leqslant p^c} \oint |\mathfrak{f}_c(\boldsymbol{\alpha};\zeta)^{2k}\mathfrak{F}_b^{\mathcal{T}}(\boldsymbol{\alpha};\eta)^{2u}|\,\mathrm{d}\boldsymbol{\alpha}.$$

A further application of Hölder's inequality shows as in (6.8) that

$$\Theta_{c,b}(X;\eta) \ll (J_{s+k}(X/M^b))^{1-k/s}(I_{b,c}(X))^{k/s},$$

and thus we find from (11.2) that

$$K_{a,b}(X) \ll M^{\frac{1}{2}k(k-1)(a+b)+k(b-a)}(J_{s+k}(X/M^b))^{1-k/s}\prod_{i=1}^{k}(I_{b,ib}(X))^{1/s}.$$

Each mean value $I_{b,ib}(X)$ may be conditioned via Lemma 5.3, and thus one deduces as in Lemma 6.3 that there exist integers $h_1,\dots,h_k$, none too large in terms of b, with the property that

$$\begin{aligned}[\![K_{a,b}(X)]\!] &\ll M^{-k/(3s)}(X/M^b)^{\eta_{s+k}}\\ &\quad + X^{\delta}(X/M^b)^{\eta_{s+k}(1-k/s)}\prod_{i=1}^{k}M^{-7h_i/4}[\![K_{b,ib+h_i}(X)]\!]^{1/s}. \qquad (11.3)\end{aligned}$$

It is (11.3) that represents the critical step in our iteration. Starting from the relation

$$X^{\eta_{s+k}-\delta} < [\![J_{s+k}(X)]\!] \ll [\![K_{0,1}(X)]\!],$$

one may apply (11.3) successively to bound $X^{\eta_{s+k}-\delta}$ in terms first of the k expressions of the shape $[\![K_{1,i+h_i}(X)]\!]$ $(1\leqslant i\leqslant k)$, then of k^2 expressions of the shape $[\![K_{b,jb+h'_j}(X)]\!]$ $(1\leqslant j\leqslant k)$, in which b takes values $i+h_i$ $(1\leqslant i\leqslant k)$, and so on. This iteration may be analysed in a manner very similar to that used in the proof of Lemma 7.2, though the complexity is now increased substantially. The important feature is the number of iterations taken before the exponents $ib+h_i$ occurring in (11.3) become large in terms of θ. In the argument of the proof of Lemma 7.2, one finds that at the nth iteration, the relevant exponents have size roughly k^n. From the relation (11.3), one obtains an explosively growing tree of chains of relations, with the exponents b_n increasing from one step to the next by a factor close to $1,2,\dots,k-1$ or k. When one considers the set of all chains, one finds that almost all possible chains have the property that the exponent b_n grows on average like $(\frac{1}{2}(k+1))^n$. In order to see this, observe that if $l_1,\dots,l_n$ are the factors at each step of one possible chain, then by the Arithmetic-Geometric Mean inequality, one has

$$l_1\cdots l_n \leqslant \Big(\frac{l_1+\cdots+l_n}{n}\Big)^n.$$

If one randomly chooses $l_1,\dots,l_n$ from $\{1,2,\dots,k\}$ with equal probability, then almost all values of $(l_1+\cdots+l_n)/n$ will be concentrated towards the mean of $\{1,2,\dots,k\}$, which is $\frac{1}{2}(k+1)$. This is a consequence of the Central Limit Theorem. In this way, one sees that the number of steps permitted

before the iteration begins to exhaust its usefulness is roughly N if we take $\theta = \frac{1}{2}((k+1)/2)^{-N-1}$ at the outset in place of $\theta = \frac{1}{2}k^{-N-1}$. Note that the latter is indeed the value that we chose for θ in Section 3 when $s = k^2$.

We are led now to a relation of similar shape to (7.14), but replaced now by

$$X^{\eta_{s+k}(1+(s/k-1)(s/k)^{N-1}\theta)} \ll X^{\eta_{s+k}+k^2}. \tag{11.4}$$

Note here that we have made use of the growth rate of the exponents ψ_n from Section 7, with scale factor s/k. Thus, when $s > \frac{1}{2}k(k+1)$, since now we have $\theta = \frac{1}{2}((k+1)/2)^{-N-1}$, we find that

$$(s/k-1)(s/k)^{N-1}\theta \gg \Big(\frac{s}{k(k+1)/2}\Big)^N,$$

which tends to infinity as N tends to infinity. In particular, on taking N sufficiently large, the relation (11.4) implies that $\eta_{s+k} = 0$.

The above heuristic shows that when $s > \frac{1}{2}k(k+1)$, then one has

$$J_{s+k}(X) \ll X^{2s+2k-\frac{1}{2}k(k+1)+\varepsilon}. \tag{11.5}$$

One might complain that this fails to prove that the relation (11.5) holds for $s = \frac{1}{2}k(k+1)$. Apart from anything else, on the face of it, the integer s needs to be a multiple of k in our treatment, so that one may need to require that $s \geqslant \frac{1}{2}k(k+3)$. But this issue may be circumvented. For this, one reinterprets the methods of this paper in the form of fractional moments of exponential sums along the lines of the author's work [40] on breaking classical convexity in Waring's problem. This was, in fact, the author's original approach to Theorem 1.1 and feasible with sufficient effort. Such would permit the proof of (11.5) with $s = \frac{1}{2}k(k+1) + \nu$ for any positive number ν. But then an application of Hölder's inequality shows that (11.5) holds with $s = \frac{1}{2}k(k+1)$ with the positive number ε bounded above by ν. Taking ν sufficiently small completes the heuristic proof.

A final word is in order concerning the value of such a heuristic argument. A more sweeping heuristic of classical nature asserts that one should expect square-root cancellation in $f_k(\boldsymbol{\alpha}; X)$ when one subtracts the expected major arc approximation, and this leads to the conjectured estimate (11.5) for $s + k \geqslant \frac{1}{2}k(k+1)$. This amounts to the assumption of very significant global rigid structure within the mean value $J_{s+k}(X)$. Our heuristic in this section also amounts to a structural assumption, but now of a rather weak congruential variety. This is, most assuredly, an unproven assumption, but a relatively modest one of local type. Thus one can say, at least, that the conjectured estimate (11.5) for $s \geqslant \frac{1}{2}k(k+1)$ now rests on only a relatively mild assumption.

References

[1] G. I. Arkhipov, The Hilbert-Kamke problem, *Izv. Akad. Nauk SSSR Ser. Mat.* **48** (1984), 3–52. MR 0733357. Zbl 0539.10016. http://dx.doi.org/10.1070/IM1985v024n01ABEH001210.

[2] G. I. Arkhipov, V. N. Chubarikov, and A. A. Karatsuba, *Trigonometric Sums in Number Theory and Analysis*, *de Gruyter Exp. Math.* **39**, Walter de Gruyter GmbH & Co. KG, Berlin, 2004. MR 2113479. Zbl 1074.11043.

[3] R. C. Baker, *Diophantine Inequalities*, *London Math. Soc. Monogr.* (*N.S.*) **1**, The Clarendon Press Oxford University Press, New York, 1986. MR 0865981. Zbl 0592.10029.

[4] B. J. Birch, Waring's problem in algebraic number fields, *Proc. Cambridge Philos. Soc.* **57** (1961), 449–459. MR 0143754. Zbl 0111.25104. http://dx.doi.org/10.1017/S0305004100035490.

[5] K. D. Boklan, The asymptotic formula in Waring's problem, *Mathematika* **41** (1994), 329–347. MR 1316613. Zbl 0815.11050. http://dx.doi.org/10.1112/S0025579300007439.

[6] K. D. Boklan and T. D. Wooley, On Weyl sums for smaller exponents, *Funct. Approx. Comment. Math.*, in press.

[7] J. Brüdern, A problem in additive number theory, *Math. Proc. Cambridge Philos. Soc.* **103** (1988), 27–33. MR 0913447. Zbl 0655.10041. http://dx.doi.org/10.1017/S0305004100064586.

[8] E. Croot and D. Hart, h-fold sums from a set with few products, *SIAM J. Discrete Math.* **24** (2010), 505–519. MR 2646099. Zbl 1221.11202. http://dx.doi.org/10.1137/090756041.

[9] K. B. Ford, New estimates for mean values of Weyl sums, *Internat. Math. Res. Notices* (1995), 155–171. MR 1321702. Zbl 0821.11050. http://dx.doi.org/10.1155/S1073792895000122.

[10] ———, Vinogradov's integral and bounds for the Riemann zeta function, *Proc. London Math. Soc.* **85** (2002), 565–633. MR 1936814. Zbl 1034.11044. http://dx.doi.org/10.1112/S0024611502013655.

[11] G. H. Hardy and J. E. Littlewood, Some problems of 'Partitio Numerorum': IV. The singular series in Waring's Problem and the value of the number $G(k)$, *Math. Z.* **12** (1922), 161–188. MR 1544511. JFM 48.0146.01. http://dx.doi.org/10.1007/BF01482074.

[12] D. R. Heath-Brown, Weyl's inequality, Hua's inequality, and Waring's problem, *J. London Math. Soc.* **38** (1988), 216–230. MR 0966294. Zbl 0619.10046. http://dx.doi.org/10.1112/jlms/s2-38.2.216.

[13] D. Hilbert, Beweis für die Darstellbarkeit der ganzen Zahlen durch eine feste Anzahl n^{ter} Potenzen (Waringsches Problem), *Math. Ann.* **67** (1909), 281–300. MR 1511530. JFM 40.0236.03. http://dx.doi.org/10.1007/BF01450405.

[14] L.-K. Hua, On Tarry's problem, *Quart. J. Math. Oxford* **9** (1938), 315–320. Zbl 0020.00501. http://dx.doi.org/10.1093/qmath/os-9.1.315.

[15] ———, Improvement of a result of Wright, *J. London Math. Soc.* **24** (1949), 157–159. MR 0030980. Zbl 0038.18104. http://dx.doi.org/10.1112/jlms/s1-24.2.157.

[16] L.-K. Hua, An improvement of Vinogradov's mean-value theorem and several applications, *Quart. J. Math., Oxford Ser.* **20** (1949), 48–61. MR 0029415. Zbl 0039.27403. http://dx.doi.org/10.1093/qmath/os-20.1.48.

[17] ———, *Additive Theory of Prime Numbers*, *Transl. Math. Monogr.* **13**, Amer. Math. Soc., Providence, R.I., 1965. MR 0194404. Zbl 0192.39304.

[18] A. A. Karatsuba, The mean value of the modulus of a trigonometric sum, *Izv. Akad. Nauk SSSR Ser. Mat.* **37** (1973), 1203–1227. MR 0337817. Zbl 0294.10025. http://dx.doi.org/10.1070/IM1973v007n06ABEH002080.

[19] Yu. V. Linnik, On Weyl's sums, *Rec. Math.* [*Mat. Sbornik*] *N.S.* **12(54)** (1943), 28–39. MR 0009776. Zbl 0063.03578. Available at http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=sm&paperid=6141&option_lang=eng.

[20] D. A. Mit′kin, An estimate for the number of summands in the Hilbert-Kamke problem, *Mat. Sb.* **129(171)** (1986), 549–577, 592. MR 0842400. Zbl 0608.10020. http://dx.doi.org/10.1070/SM1987v057n02ABEH003087.

[21] ———, An estimate for the number of summands in the Hilbert-Kamke problem. II, *Mat. Sb.* **132(174)** (1987), 345–351, 444–445. MR 0889596. Zbl 0619.10013. http://dx.doi.org/10.1070/SM1988v060n02ABEH003172.

[22] S. T. Parsell, A generalization of Vinogradov's mean value theorem, *Proc. London Math. Soc.* **91** (2005), 1–32. MR 2149529. Zbl 1119.11024. http://dx.doi.org/10.1112/S002461150501525X.

[23] ———, On the Bombieri–Korobov estimate for Weyl sums, *Acta Arith.* **138** (2009), 363–372. MR 2534142. Zbl 05615154. http://dx.doi.org/10.4064/aa138-4-7.

[24] O. Robert and P. Sargos, Un théorème de moyenne pour les sommes d'exponentielles. Application à l'inégalité de Weyl, *Publ. Inst. Math.* (*Beograd*) **67(81)** (2000), 14–30. MR 1761299. Zbl 1006.11046.

[25] W. M. Schmidt, The density of integer points on homogeneous varieties, *Acta Math.* **154** (1985), 243–296. MR 0781588. Zbl 0561.10010. http://dx.doi.org/10.1007/BF02392473.

[26] S. B. Stechkin, Mean values of the modulus of a trigonometric sum, *Trudy Mat. Inst. Steklov.* **134** (1975), 283–309, 411. MR 0396431. Zbl 0319.10045. Available at http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=tm&paperid=2719&option_lang=eng.

[27] A. V. Ustinov, On the number of summands in the asymptotic formula for the number of solutions of the Waring equation, *Mat. Zametki* **64** (1998), 285–296. MR 1681004. Zbl 0923.11138. http://dx.doi.org/10.1007/BF02310311.

[28] R. C. Vaughan, On Waring's problem for cubes, *J. Reine Angew. Math.* **365** (1986), 122–170. MR 0826156. Zbl 0574.10046. http://dx.doi.org/10.1515/crll.1986.365.122.

[29] ———, On Waring's problem for smaller exponents. II, *Mathematika* **33** (1986), 6–22. MR 0859494. Zbl 0601.10037. http://dx.doi.org/10.1112/S0025579300013838.

[30] ———, *The Hardy-Littlewood Method*, second ed., *Cambridge Tracts in Math.* **125**, Cambridge Univ. Press, Cambridge, 1997. MR 1435742. Zbl 0868.11046.

[31] R. C. Vaughan and T. D. Wooley, A special case of Vinogradov's mean value theorem, *Acta Arith.* **79** (1997), 193–204. MR 1438823. Zbl 0887.11042.

[32] I. M. Vinogradov, New estimates for Weyl sums, *Dokl. Akad. Nauk SSSR* **8** (1935), 195–198.

[33] ———, The method of trigonometrical sums in the theory of numbers, *Trav. Inst. Math. Stekloff* **23** (1947), 109 pp. MR 0029417. Zbl 0041.37002.

[34] ———, A new estimate of the function $\zeta(1 + it)$, *Izv. Akad. Nauk SSSR. Ser. Mat.* **22** (1958), 161–164. MR 0103861. Zbl 0097.26302. Available at http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=im&paperid=3962&option_lang=eng.

[35] A. Z. Walfisz, *Weylsche Exponentialsummen in der neueren Zahlentheorie*, *Math. Forsch.* **15**, Berlin, 1963. Zbl 0146.06003.

[36] T. D. Wooley, Large improvements in Waring's problem, *Ann. of Math.* **135** (1992), 131–164. MR 1147960. Zbl 0754.11026. http://dx.doi.org/10.2307/2946566.

[37] ———, On Vinogradov's mean value theorem, *Mathematika* **39** (1992), 379–399. MR 1203293. Zbl 0769.11036. http://dx.doi.org/10.1112/S0025579300015102.

[38] ———, On Vinogradov's mean value theorem, II, *Michigan Math. J.* **40** (1993), 175–180. MR 1214062. Zbl 0805.11072. http://dx.doi.org/10.1307/mmj/1029004681.

[39] ———, Quasi-diagonal behaviour in certain mean value theorems of additive number theory, *J. Amer. Math. Soc.* **7** (1994), 221–245. MR 1224595. Zbl 0786.11053. http://dx.doi.org/10.2307/2152728.

[40] ———, Breaking classical convexity in Waring's problem: sums of cubes and quasi-diagonal behaviour, *Invent. Math.* **122** (1995), 421–451. MR 1359599. Zbl 0851.11055. http://dx.doi.org/10.1007/BF01231451.

[41] ———, New estimates for Weyl sums, *Quart. J. Math. Oxford Ser.* **46** (1995), 119–127. MR 1326136. Zbl 0855.11043. http://dx.doi.org/10.1093/qmath/46.1.119.

[42] ———, Some remarks on Vinogradov's mean value theorem and Tarry's problem, *Monatsh. Math.* **122** (1996), 265–273. MR 1414861. Zbl 0881.11043. http://dx.doi.org/10.1007/BF01320189.

[43] ———, The asymptotic formula in Waring's problem, *Internat. Math. Res. Notices* **(2012),** no. 7, 1485–1504. http://dx.doi.org/10.1093/imrn/rnr074.

[44] E. M. Wright, The Prouhet-Lehmer problem, *J. London Math. Soc.* **23** (1948), 279–285. MR 0028858. Zbl 0033.35102. http://dx.doi.org/10.1112/jlms/s1-23.4.279.

(Received: December 3, 2010)
(Revised: July 11, 2011)

School of Mathematics, University of Bristol, Bristol, UK
E-mail: matdw@bristol.ac.uk