DISCRETE FOURIER RESTRICTION VIA EFFICIENT CONGRUENCING: BASIC PRINCIPLES

TREVOR D. WOOLEY

Abstract. We show that whenever $s > k(k+1)$, then for any complex sequence $(\mathfrak{a}_n)_{n\in\mathbb{Z}}$, one has

$$\int_{[0,1)^k} \Bigl| \sum_{|n|\leqslant N} \mathfrak{a}_n e(\alpha_1 n + \ldots + \alpha_k n^k) \Bigr|^{2s} \mathrm{d}\boldsymbol{\alpha} \ll N^{s-k(k+1)/2} \Bigl(\sum_{|n|\leqslant N} |\mathfrak{a}_n|^2 \Bigr)^s .$$

Bounds for the constant in the associated periodic Strichartz inequality from L^{2s} to l^2 of the conjectured order of magnitude follow, and likewise for the constant in the discrete Fourier restriction problem from l^2 to $L^{s'}$, where $s' = 2s/(2s-1)$. These bounds are obtained by generalising the efficient congruencing method from Vinogradov's mean value theorem to the present setting, introducing tools of wider application into the subject.

1. Introduction

Our goal in this paper is to introduce a flexible new method for analysing a wide class of Fourier restriction problems, illustrating our approach in this first instance with a model example. In order to set the scene, consider a natural number k and a function $g : (\mathbb{R}/\mathbb{Z})^k \to \mathbb{C}$ having an associated Fourier series

$$\tilde{g}(\alpha_1, \ldots, \alpha_k) = \sum_{\mathbf{n}\in\mathbb{Z}^k} \hat{g}(n_1, \ldots, n_k) e(n_1\alpha_1 + \ldots + n_k\alpha_k). \tag{1.1}$$

Here, we permit the Fourier coefficients $\hat{g}(\mathbf{n})$ to be arbitrary complex numbers, and as usual write $e(z) = e^{2\pi i z}$. Beginning with work of Stein (see [17], and [2] and [18] for more recent broader context), there is by now an extensive body of research concerning the norms of operators restricting such Fourier series to integral points $\mathbf{n}$ lying on manifolds of various dimensions. Thus, for example, in work concerning the non-linear Schrödinger and KdV equations, Bourgain [4, 5] has considered the situation with $k = 2$ and the restriction $\mathbf{n} = (n, n^l)$ to the Fourier series

$$\mathcal{R}g = \sum_{n\in\mathbb{Z}} \hat{g}(n, n^l) e(n\alpha_1 + n^l\alpha_2) \quad (l = 2, 3),$$

as well as higher dimensional analogues (see [10, 11] for recent work on these problems). Such results have also found recent application in additive combinatorics in work concerning the solutions of translation invariant equations with variables restricted to dense subsets of the integers (see [8, 14, 15]).

2010 *Mathematics Subject Classification.* 42B05, 11L07, 42A16.

Key words and phrases. Fourier series, exponential sums, restriction theory, Strichartz inequalities.

We now focus on the example central to this paper, namely that in which the integral points $\mathbf{n}$ are restricted to lie on the curve $(n, n^2, \dots, n^k)$. When $(\mathfrak{a}_n)_{n\in\mathbb{Z}}$ is a sequence of complex numbers, write

$$f_{\mathfrak{a}}(\boldsymbol{\alpha}; N) = \sum_{|n|\leqslant N} \mathfrak{a}_n e(\alpha_1 n + \alpha_2 n^2 + \dots + \alpha_k n^k). \tag{1.2}$$

Our goal is to obtain the periodic Strichartz inequality

$$\|f_{\mathfrak{a}}(\boldsymbol{\alpha}; N)\|_{L^p} \leqslant K_{p,N} \|\mathfrak{a}_n\|_{l^2} \quad (p \geqslant 2), \tag{1.3}$$

with the sharpest attainable constant $K_{p,N}$, uniformly in $(\mathfrak{a}_n)$. By duality, this problem is related to the discrete restriction problem of obtaining the sharpest attainable constant $A_{p,N}$ for which

$$\sum_{|n|\leqslant N} |\hat{g}(n, n^2, \dots, n^k)|^2 \leqslant A_{p,N} \|g\|_{p'}^2 \quad (p \geqslant 2), \tag{1.4}$$

where $g : (\mathbb{R}/\mathbb{Z})^k \to \mathbb{C}$, and $p' = p/(p-1)$. Indeed, one has $K_{p,N} \sim A_{p,N}^{1/2}$.

By adapting the efficient congruencing method introduced in work [21] of the author associated with Vinogradov's mean value theorem, we obtain in §§6 and 7 the following conclusion.

Theorem 1.1. *Suppose that $k \geqslant 2$ and $s \geqslant k(k+1)$. Then for any $\varepsilon > 0$, and any complex sequence $(\mathfrak{a}_n)_{n\in\mathbb{Z}}$, one has*[1]

$$\oint |f_{\mathfrak{a}}(\boldsymbol{\alpha}; N)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \ll N^{s-k(k+1)/2+\varepsilon} \Big(\sum_{|n|\leqslant N} |\mathfrak{a}_n|^2 \Big)^s. \tag{1.5}$$

Moreover, when $s > k(k+1)$, one may take $\varepsilon = 0$.

By way of comparison, we note that when $k = 2$, Bourgain [4] has obtained the estimate

$$\oint |f_{\mathfrak{a}}(\boldsymbol{\alpha}; N)|^{s} \, \mathrm{d}\boldsymbol{\alpha} \ll N^{\varepsilon}(1 + N^{s-3}) \Big(\sum_{|n|\leqslant N} |\mathfrak{a}_n|^2 \Big)^s,$$

in which the factor N^ε may be deleted whenever $s \neq 6$. Indeed, Bourgain shows that when $s = 6$ this factor may be replaced by one of the shape $\exp(c \log N / \log\log N)$, and that it cannot be deleted. Also, forthcoming work of Kevin Hughes [13] delivers a similar conclusion to that of Theorem 1.1, though only for values of s roughly twice as large as demanded by our new theorem.

Recall (1.3) and (1.4). In §8 we show that the conclusion of Theorem 1.1 yields the following corollary providing bounds for $K_{p,N}$ and $A_{p,N}$.

---

[1]We employ the convention that whenever $G : [0,1)^k \to \mathbb{C}$ is integrable, then $\oint G(\boldsymbol{\alpha}) \, \mathrm{d}\boldsymbol{\alpha} = \int_{[0,1)^k} G(\boldsymbol{\alpha}) \, \mathrm{d}\boldsymbol{\alpha}$. Moreover, constants implicit in Vinogradov's notation $\ll$ and $\gg$ may depend on s, k and ε.

Corollary 1.2. *When $k \geqslant 2$, $p \geqslant 2k(k+1)$ and $\varepsilon > 0$, one has*

$$K_{p,N} \ll N^{(1-\theta)/2+\varepsilon} \quad \text{and} \quad A_{p,N} \ll N^{1-\theta+\varepsilon},$$

where we write $\theta = k(k+1)/p$. Moreover, provided that $p > 2k(k+1)$, one may take $\varepsilon = 0$.

Finally, a straightforward argument in §8 conveys us from the estimates provided by Theorem 1.1 to the bounds recorded in the following corollary.

Corollary 1.3. *Suppose that $t \geqslant 1$ and that $k_1, \ldots, k_t$ are positive integers with $1 \leqslant k_1 < k_2 < \ldots < k_t = k$. Let $s \geqslant k(k+1)$, and write $K = k_1 + \ldots + k_t$. Then for any $\varepsilon > 0$, and any complex sequence $(\mathfrak{a}_n)$, one has*

$$\oint \Bigl| \sum_{|n| \leqslant N} \mathfrak{a}_n e(\alpha_1 n^{k_1} + \ldots + \alpha_t n^{k_t}) \Bigr|^{2s} \mathrm{d}\boldsymbol{\alpha} \ll N^{s-K+\varepsilon} \Bigl(\sum_{|n| \leqslant N} |\mathfrak{a}_n|^2 \Bigr)^{s}.$$

Moreover, when $s > k(k+1)$, one may take $\varepsilon = 0$.

Bounds of the shape supplied by Theorem 1.1 are closely related to those available in Vinogradov's mean value theorem, which corresponds to the special case in which $(\mathfrak{a}_n) = (1)$. In the latter circumstances, one has the lower bound

$$\oint \Bigl| \sum_{|n| \leqslant N} e(\alpha_1 n + \ldots + \alpha_k n^k) \Bigr|^{2s} \mathrm{d}\boldsymbol{\alpha} \gg N^s + N^{2s-k(k+1)/2}. \tag{1.6}$$

It is therefore of interest to determine the least number $s_0 = s_0(k)$ for which, for any $\varepsilon > 0$, one has the corresponding upper bound

$$\oint \Bigl| \sum_{|n| \leqslant N} e(\alpha_1 n + \ldots + \alpha_k n^k) \Bigr|^{2s} \mathrm{d}\boldsymbol{\alpha} \ll N^{2s-k(k+1)/2+\varepsilon}.$$

The classical work of Vinogradov [20] and Hua [12] shows that one may take $s_0(k) \leqslant 3k^2(\log k + O(\log\log k))$ for large k (see Vaughan [19, Theorem 7.4], for example). Very recently, with the arrival of the efficient congruencing method, such conclusions have become available for all $k \geqslant 2$ with $s_0(k) \leqslant k(k+1)$ (see Wooley [21, Theorem 1.1]). The very latest refinements of such work [25, Theorem 1.2] show that one may even take $s_0(k) \leqslant k(k-1)$ whenever $k \geqslant 3$.

Following the discussion of the last paragraph, we are equipped to describe some consequences of forthcoming work [13] of Kevin Hughes. This delivers a conclusion of the shape (1.5) provided that $s \geqslant 2s_0(k)$, showing also that one may take $\varepsilon = 0$ for $s > 2s_0(k)$. One can interpret Hughes' method as bounding mean values of $f_{\mathfrak{a}}(\boldsymbol{\alpha}; N)$ in terms of corresponding mean values of classical Vinogradov type with half as many underlying variables. This work demands essentially twice as many variables as are required in our Theorem 1.1. In this paper, we work directly with mean values of $f_{\mathfrak{a}}(\boldsymbol{\alpha}; N)$, avoiding any reference to the classical version of Vinogradov's mean value theorem, and avoiding the losses inherent in previous approaches.

When $s > k(k+1)$, in the special case of the sequence $(\mathfrak{a}_n) = (1)$, the lower bound (1.6) confirms that the upper bound (1.5) of Theorem 1.1 is sharp (in which case one may take $\varepsilon = 0$). Indeed, by reference to this same special sequence, one may formulate the following conjecture.

Conjecture 1.4 (Main Conjecture)**.** *Suppose that $k \geqslant 1$. Then for any $\varepsilon > 0$, and any complex sequence $(\mathfrak{a}_n)$, one has*

$$\oint |f_{\mathfrak{a}}(\boldsymbol{\alpha}; N)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \ll N^{\varepsilon} \left(1 + N^{s-k(k+1)/2}\right) \Bigl(\sum_{|n| \leqslant N} |\mathfrak{a}_n|^2 \Bigr)^{s}.$$

By orthogonality, the mean value on the left hand side here is a weighted count of the number of integral solutions of the Diophantine system

$$\sum_{i=1}^{s} x_i^j = \sum_{i=1}^{s} y_i^j \quad (1 \leqslant j \leqslant k), \tag{1.7}$$

with $|x_i|, |y_i| \leqslant N$. When $1 \leqslant s \leqslant k$, it follows from Newton's formulae concerning roots of polynomials that $\{x_1, \ldots, x_s\} = \{y_1, \ldots, y_s\}$, and thus

$$\oint |f_{\mathfrak{a}}(\boldsymbol{\alpha}; N)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \ll \Bigl(\sum_{|n| \leqslant N} |\mathfrak{a}_n|^2 \Bigr)^{s}.$$

Thus, Conjecture 1.4 holds when $1 \leqslant s \leqslant k$ for essentially trivial reasons. With work, this line of reasoning can be extended to cover the case $s = k + 1$. We note that Bourgain and Demeter [6] have confirmed this conjecture in the longer range $1 \leqslant s \leqslant 2k - 1$. By a substantial elaboration of the ideas of this paper, which we intend to pursue within a more general framework in a future memoir, we are able to extend this range substantially so as to confirm Conjecture 1.4 in the interval $1 \leqslant s \leqslant D(k)$, where $D(4) = 8$, $D(5) = 10$, ..., and $D(k) = \frac{1}{2}k(k+1) - (\frac{1}{3} + o(1))k$. Note that the confirmation of this conjecture for $1 \leqslant s \leqslant k(k+1)/2$ would suffice to confirm it for all $s \geqslant 1$.

We have restricted ourselves in this paper to the relatively more straightforward proof of Theorem 1.1 so as to make the ideas underlying this generalisation of the efficient congruencing method more transparent. We hope, in this way, to permit other workers more easily to consider the use of such methods in applications farther afield.

Our basic strategy in proving Theorem 1.1 is to adapt to the present setting the efficient congruencing method introduced by the author in the context of Vinogradov's mean value theorem (see [21]). Several complications must be surmounted in such a plan of attack. First, the presence of arbitrary complex coefficients $\mathfrak{a}_n$ implies that the strong translation-dilation invariance present in the setting of Vinogradov's mean value theorem is absent. However, we are able to normalise the exponential sums $f_{\mathfrak{a}}(\boldsymbol{\alpha}; N)$ by a factor $\left(\sum_{|n| \leqslant N} |\mathfrak{a}_n|^2\right)^{-1/2}$ so as to achieve scale invariance, and subsequently consider mean values associated with extremal sequences. This initial preparation is discussed in §2, and recovers a sufficient approximation to translation invariance that subsequent operations are not impeded.

We next impose a congruential condition, modulo an auxiliary prime number ϖ, on the summands underlying the mean value on the left hand side of

(1.5). By employing the weak translation-dilation invariance present in the underlying Diophantine system, as discussed in §3, we find that a subset of the summands are subject to a strong congruence condition of the shape

$$\sum_{i=1}^k x_i^j \equiv \sum_{i=1}^k y_i^j \pmod{\varpi^j} \quad (1 \leqslant j \leqslant k).$$

By incorporating a non-singularity condition *en passant* in §4, one extracts in §5 the condition $x_i \equiv y_i \pmod{\varpi^k}$ on the underlying variables. In this way, the original mean value is bounded above by a new mean value, both old and new containing complex weights, and the new one subject to powerful congruence constraints. By appropriate application of Hölder's inequality, this new mean value is bounded above by a further mean value encoding still stronger congruence constraints.

As in the efficient congruencing method for Vinogradov's mean value theorem, one now seeks to utilise this congruence concentration argument. In essence, if the original mean value is assumed to be significantly larger than conjectured, then one can show that a certain auxiliary mean value is significantly larger still by comparison to its expected size. By iterating, one ultimately arrives at a mean value shown to be so much larger than expected, that it exceeds even a trivial estimate for its magnitude. Thus one contradicts the initial assumption, and one is forced to conclude that the original mean value is of approximately the same magnitude as anticipated. This process is discussed in §6. The additional complication associated with this approach concerns the complex weights $\mathfrak{a}_n$. For this reason, one must incorporate extra averaging by comparison with the earlier efficient congruencing approach.

One last matter deserves attention, namely that of the claim to the effect that one may take $\varepsilon = 0$ in the conclusion of Theorem 1.1 when $s > k(k+1)$. Such a conclusion first became available in the special case $k = 2$ and $s > 3$ in the work of Bourgain [4]. This approach has been generalised in forthcoming work of Kevin Hughes [13][2], so that whenever one has a bound of the shape

$$\oint |f_{\mathfrak{a}}(\boldsymbol{\alpha}; N)|^p \, \mathrm{d}\boldsymbol{\alpha} \ll N^{(p-k(k+1))/2+\varepsilon} \Biggl(\sum_{|n| \leqslant N} |\mathfrak{a}_n|^2 \Biggr)^{p/2},$$

valid for $p \geqslant p_0(k)$, then the same conclusion holds with $\varepsilon = 0$ provided that $p > p_0(k)$. Although we could apply Hughes' ε-removal lemma, we have opted in §7 instead for a cheap treatment of slightly less generality in order that our account be self-contained. This approach is based on the Keil-Zhao device (see [14] and [26]), and we hope that it may be of independent interest.

It will be apparent to experts that the ideas introduced in this paper to surmount the difficulties associated with the complex weights $(\mathfrak{a}_n)$ are very flexible. Indeed, it seems that there is a metamathematical principle available that, when given a conventional unweighted mean value estimate established

[2]Very recently, the author has been informed by Kevin Henriot that he has independently obtained such an ε-removal lemma, and that this will appear in his forthcoming paper [9].

by a variant of the efficient congruencing method, delivers the corresponding estimate equipped with arbitrary complex weights through the methods of this paper. We discuss some of the immediate consequences of this principle in §9.

The author is very grateful to Kevin Hughes for an inspiring seminar and subsequent conversations that provided motivation for this paper.

2. The infrastructure for efficient congruencing

We initiate our discussion of the proof of Theorem 1.1 by introducing the components and basic notation required to assemble the weighted efficient congruencing iteration. Although analogous to that of our corresponding work [21] concerning Vinogradov's mean value theorem, we are forced to deviate significantly from our earlier path.

Let k be a fixed integer with $k \geqslant 2$, consider a complex sequence $(\mathfrak{a}_n)$, and recall the exponential sum $f_{\mathfrak{a}}(\boldsymbol{\alpha}; N)$ defined in (1.2). Our first task is to replace the mean value central to Theorem 1.1 by a modification less intimately dependent on the sequence $(\mathfrak{a}_n)$. In this context, we remark that in the main thrust of our argument, we restrict attention to sequences other than the zero sequence $(\mathfrak{a}_n) = (0)$. Define $\rho(\mathfrak{a}; N)$ to be 1 when $(\mathfrak{a}_n) = (0)$, and otherwise by taking

$$\rho(\mathfrak{a}; N) = \Bigg(\sum_{|n| \leqslant N} |\mathfrak{a}_n|^2 \Bigg)^{1/2}. \tag{2.1}$$

We define the normalised exponential sum

$$\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha}; N) = \rho(\mathfrak{a}; N)^{-1} \sum_{|n| \leqslant N} \mathfrak{a}_n e(\psi(n; \boldsymbol{\alpha})), \tag{2.2}$$

in which we put

$$\psi(n; \boldsymbol{\alpha}) = \alpha_k n^k + \ldots + \alpha_1 n. \tag{2.3}$$

Then, when $s > 0$, we define the mean value

$$U_{s,k}(N; \mathfrak{a}) = \oint |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha}; N)|^{2s} \, \mathrm{d}\boldsymbol{\alpha}. \tag{2.4}$$

A comparison of (1.2) and (2.2) reveals that

$$\oint |f_{\mathfrak{a}}(\boldsymbol{\alpha}; N)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \ll \rho(\mathfrak{a}; N)^{2s} U_{s,k}(N; \mathfrak{a}) = U_{s,k}(N; \mathfrak{a}) \Bigg(\sum_{|n| \leqslant N} |\mathfrak{a}_n|^2 \Bigg)^{s}.$$

The proof of Theorem 1.1 will therefore be accomplished by establishing that, whenever $s \geqslant k(k+1)$, then for any $\varepsilon > 0$ and any complex sequence $(\mathfrak{a}_n)$, one has

$$U_{s,k}(N; \mathfrak{a}) \ll N^{s - k(k+1)/2 + \varepsilon},$$

and further that one may take $\varepsilon = 0$ when $s > k(k+1)$. Notice, in this context, that when $(\mathfrak{a}_n) = (0)$, then $U_{s,k}(N; \mathfrak{a}) = 0$, so that we are entitled to ignore the zero sequence in subsequent discussion.

By applying Cauchy's inequality to (2.2), one obtains the bound

$$|\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};N)| \leqslant \rho(\mathfrak{a};N)^{-1}\Big(\sum_{|n|\leqslant N}|\mathfrak{a}_n|^2\Big)^{1/2}(2N+1)^{1/2},$$

and thus it follows from (2.1) that whenever $N\geqslant 1$, one has

$$|\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};N)| \leqslant 2N^{1/2}, \tag{2.5}$$

uniformly in $(\mathfrak{a}_n)$ and $\boldsymbol{\alpha}$. In particular, we infer from the definition (2.4) that $U_{s,k}(N;\mathfrak{a})\ll N^s$, uniformly in $(\mathfrak{a}_n)$. Further, by taking $(\mathfrak{a}_n)=(1)$ and applying orthogonality, one finds from (1.2) that when $s\in\mathbb{N}$, the mean value $\oint|f_{\mathfrak{a}}(\boldsymbol{\alpha};N)|^{2s}\,\mathrm{d}\boldsymbol{\alpha}$ counts the number of integral solutions of the system of equations (1.7) with $|x_i|,|y_i|\leqslant N$. The contribution of the diagonal solutions $x_i=y_i$ $(1\leqslant i\leqslant s)$ ensures in such circumstances that

$$\rho(\mathfrak{a};N)^{-2s}\oint|f_{\mathfrak{a}}(\boldsymbol{\alpha};N)|^{2s}\,\mathrm{d}\boldsymbol{\alpha}\gg(N^{1/2})^{-2s}N^s=1.$$

We therefore deduce that when $(\mathfrak{a}_n)=(1)$, then one has $U_{s,k}(N;\mathfrak{a})\gg 1$.

The definition of $U_{s,k}(N;\mathfrak{a})$ ensures that it is non-negative for all N and $(\mathfrak{a}_n)$. In addition, one sees from (2.2) that $\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};N)$, and hence also $U_{s,k}(N;\mathfrak{a})$, is scale invariant with respect to $(\mathfrak{a}_n)$, meaning that $U_{s,k}(N;\gamma\mathfrak{a})=U_{s,k}(N;\mathfrak{a})$ for any $\gamma>0$. We may therefore suppose without loss of generality that, for any fixed value of N, one has $|\mathfrak{a}_n|\leqslant 1$ for $|n|\leqslant N$. Write $\mathbb{D}$ for the unit disc $\{z\in\mathbb{C}:|z|\leqslant 1\}$. Then it follows that for each fixed $N\geqslant 1$, one has

$$0\leqslant \sup_{\substack{\mathfrak{a}_n\in\mathbb{D}\ (|n|\leqslant N)\\ (\mathfrak{a}_n)\neq(0)}}\frac{\log U_{s,k}(N;\mathfrak{a})}{\log N}\leqslant s.$$

Thus, when $s\in\mathbb{N}$, we may define the quantity

$$\lambda_{s,k}=\limsup_{N\to\infty}\sup_{\substack{\mathfrak{a}_n\in\mathbb{D}\ (|n|\leqslant N)\\ (\mathfrak{a}_n)\neq(0)}}\frac{\log U_{s,k}(N;\mathfrak{a})}{\log N}. \tag{2.6}$$

We make one further simplification before proceeding further, observing that there is no loss of generality in restricting the supremum in (2.6) to be taken over complex sequences $(\mathfrak{a}_n)$ all of whose terms are real and positive. For given a complex sequence $(\mathfrak{a}_n)$, put $\mathfrak{b}_n=|\mathfrak{a}_n|$ for each n. Then it follows from (2.1) that $\rho(\mathfrak{b};N)=\rho(\mathfrak{a};N)$. Moreover, by orthogonality, the mean value $U_{s,k}(N;\mathfrak{a})$ counts the number of integral solutions of the system of equations (1.7) with $|x_i|,|y_i|\leqslant N$, each solution $\mathbf{x},\mathbf{y}$ being counted with weight

$$\rho(\mathfrak{a};N)^{-2s}\mathfrak{a}_{x_1}\dots\mathfrak{a}_{x_s}\overline{\mathfrak{a}}_{y_1}\dots\overline{\mathfrak{a}}_{y_s}.$$

This weight has absolute value equal to

$$\rho(\mathfrak{b};N)^{-2s}\mathfrak{b}_{x_1}\dots\mathfrak{b}_{x_s}\overline{\mathfrak{b}}_{y_1}\dots\overline{\mathfrak{b}}_{y_s}, \tag{2.7}$$

and so by reversing the orthogonality argument, we find that

$$U_{s,k}(N;\mathfrak{a})\leqslant U_{s,k}(N;\mathfrak{b}).$$

Moreover, should any sequence element $\mathfrak{b}_n$ be equal to 0, then the weight (2.7) changes by a quantity lying in the interval $[0,\omega)$ when we substitute $\mathfrak{b}_n=\omega>0$. By repeating this process for each sequence element, and considering the limit as $\omega\to 0$, it is apparent that

$$\sup_{\substack{\mathfrak{a}_n\in\mathbb{D}\ (|n|\leqslant N)\\ (\mathfrak{a}_n)\neq(0)}}\frac{\log U_{s,k}(N;\mathfrak{a})}{\log N}=\sup_{\mathfrak{b}_n\in(0,1]\ (|n|\leqslant N)}\frac{\log U_{s,k}(N;\mathfrak{b})}{\log N},$$

and so we are at liberty to restrict attention throughout the ensuing discussion to sequences of positive numbers. In particular, we may replace (2.6) by the equivalent definition

$$\lambda_{s,k}=\limsup_{N\to\infty}\sup_{\mathfrak{a}_n\in(0,1]\ (|n|\leqslant N)}\frac{\log U_{s,k}(N;\mathfrak{a})}{\log N}.$$

In what follows, we fix, once and for all, the integer $k\geqslant 2$, and henceforth omit its explicit mention in our notation. Furthermore, we consider a natural number u with $u\geqslant k$, and we put $s=uk$. Our method is iterative, with N iterations, and we suppose this integer to be sufficiently large in terms of s and k. We then put

$$\theta=(Nk)^{-3N}\quad\text{and}\quad\delta=(N^2s)^{-3N},\tag{2.8}$$

so that, in particular, the parameter δ is small compared to θ. These quantities will shortly make an appearance that explains their role in the argument. Rather than discuss moments of order $2s$, it is more convenient to consider $U_{s+k}(X;\mathfrak{a})$. The definition of $\lambda=\lambda_{s+k,k}$ in (2.6) now ensures that there exists a sequence $(X_n)_{n=1}^\infty$ with $\lim\limits_{n\to\infty}X_n=+\infty$ such that the following two statements hold whenever m is sufficiently large. First, for some sequence $(\mathfrak{a}_n)$ of real numbers with $\mathfrak{a}_n\in(0,1]$, one has that

$$U_{s+k}(X_m;\mathfrak{a})>X_m^{\lambda-\delta}.\tag{2.9}$$

Second, whenever $X_m^{1/2}\leqslant Y\leqslant X_m$, then for all non-zero complex sequences $(\mathfrak{a}'_n)$, one has

$$U_{s+k}(Y;\mathfrak{a}')<Y^{\lambda+\delta}.\tag{2.10}$$

We focus henceforth on a fixed element $X=X_m$ of the sequence (X_n), which we may assume to be sufficiently large in terms of s, k, N and δ. We also fix a real sequence $(\mathfrak{a}_n)$ with $\mathfrak{a}_n\in(0,1]$, satisfying (2.9). It is convenient in the remainder of §§2–6 to abbreviate $U_{s+k}(X;\mathfrak{a})$ to $U(X;\mathfrak{a})$, or even to $U(X)$.

We next recall some standard notational conventions. The letter ε denotes a sufficiently small positive number. We think of the basic parameter as being X, a large real number depending at most on ε and the other ambient parameters as indicated. Whenever ε appears in a statement, we assert that the statement holds for each $\varepsilon>0$. As usual, we write $\lfloor\psi\rfloor$ to denote the largest integer no larger than ψ, and $\lceil\psi\rceil$ to denote the least integer no smaller than ψ. We make sweeping use of vector notation. Thus, with t implied from the environment at hand, we write $\mathbf{z}\equiv\mathbf{w}\pmod{\varpi}$ to denote that $z_i\equiv w_i\pmod{\varpi}$ $(1\leqslant i\leqslant t)$, or $\mathbf{z}\equiv\xi\pmod{\varpi}$ to denote that $z_i\equiv\xi\pmod{\varpi}$ $(1\leqslant i\leqslant t)$.

Congruences to prime power moduli lie at the heart of our argument. Put $M = X^\theta$, and note that $X^\delta < M^{1/N}$. Let ϖ be a fixed prime number with $M < \varpi \leqslant 2M$ to be chosen in due course. That such a prime exists is a consequence of the Prime Number Theorem. When c and ξ are non-negative integers, we define $\rho_c(\xi) = \rho_c(\xi; \mathfrak{a})$ by putting

$$\rho_c(\xi) = \Bigg(\sum_{\substack{|n| \leqslant X \\ n \equiv \xi \ (\mathrm{mod}\ \varpi^c)}} |\mathfrak{a}_n|^2 \Bigg)^{1/2}. \tag{2.11}$$

Note that, in terms of our earlier notation, one has

$$\rho_0(1) = \rho(\mathfrak{a}; X).$$

Moreover, one has the trivial relation

$$\rho_0(1)^2 = \sum_{\xi=1}^{\varpi^c} \rho_c(\xi)^2.$$

In view of our assumption that $\mathfrak{a}_n \in (0, 1]$ for each n, it is apparent that one has $\rho_c(\xi) > 0$ whenever $1 \leqslant \xi < X$.

Recalling the notation (2.3), we introduce the normalised exponential sum

$$\mathfrak{f}_c(\boldsymbol{\alpha}; \xi) = \rho_c(\xi)^{-1} \sum_{\substack{|n| \leqslant X \\ n \equiv \xi \ (\mathrm{mod}\ \varpi^c)}} \mathfrak{a}_n e(\psi(n; \boldsymbol{\alpha})). \tag{2.12}$$

We find it necessary to consider *well-conditioned* k-tuples of integers belonging to distinct congruence classes modulo a suitable power of ϖ. Denote by $\Xi_c(\xi)$ the set of k-tuples $(\xi_1, \ldots, \xi_k)$, with

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

and satisfying the property that $\xi_i \equiv \xi_j \ (\mathrm{mod}\ \varpi^{c+1})$ for no suffices i and j with $1 \leqslant i < j \leqslant k$. We then put

$$\mathfrak{F}_c(\boldsymbol{\alpha}; \xi) = \rho_c(\xi)^{-k} \sum_{\boldsymbol{\xi} \in \Xi_c(\xi)} \prod_{i=1}^{k} \rho_{c+1}(\xi_i) \mathfrak{f}_{c+1}(\boldsymbol{\alpha}; \xi_i). \tag{2.13}$$

When a and b are non-negative integers, we define

$$I_{a,b}(X; \xi, \eta) = \oint |\mathfrak{F}_a(\boldsymbol{\alpha}; \xi)^2 \mathfrak{f}_b(\boldsymbol{\alpha}; \eta)^{2s}| \, \mathrm{d}\boldsymbol{\alpha}, \tag{2.14}$$

$$K_{a,b}(X; \xi, \eta) = \oint |\mathfrak{F}_a(\boldsymbol{\alpha}; \xi)^2 \mathfrak{F}_b(\boldsymbol{\alpha}; \eta)^{2u}| \, \mathrm{d}\boldsymbol{\alpha}, \tag{2.15}$$

and then put

$$I_{a,b}(X)=\rho_0(1)^{-4}\sum_{\xi=1}^{\varpi^a}\sum_{\eta=1}^{\varpi^b}\rho_a(\xi)^2\rho_b(\eta)^2I_{a,b}(X;\xi,\eta), \tag{2.16}$$

$$K_{a,b}(X)=\rho_0(1)^{-4}\sum_{\xi=1}^{\varpi^a}\sum_{\eta=1}^{\varpi^b}\rho_a(\xi)^2\rho_b(\eta)^2K_{a,b}(X;\xi,\eta). \tag{2.17}$$

By orthogonality, the mean value $I_{a,b}(X;\xi,\eta)$ counts the number of integral solutions of the system

$$\sum_{i=1}^k(x_i^j-y_i^j)=\sum_{l=1}^s(v_l^j-w_l^j)\quad (1\leqslant j\leqslant k), \tag{2.18}$$

with

$$|\mathbf{x}|,|\mathbf{y}|,|\mathbf{v}|,|\mathbf{w}|\leqslant X,\quad \mathbf{x},\mathbf{y}\in \Xi_a(\xi)\ (\mathrm{mod}\ \varpi^{a+1}), \tag{2.19}$$

and $\mathbf{v}\equiv\mathbf{w}\equiv\eta\ (\mathrm{mod}\ \varpi^b)$, each solution being counted with weight

$$\rho_a(\xi)^{-2k}\rho_b(\eta)^{-2s}\Bigl(\prod_{i=1}^k\mathfrak{a}_{x_i}\mathfrak{a}_{y_i}\Bigr)\Bigl(\prod_{l=1}^s\mathfrak{a}_{v_l}\mathfrak{a}_{w_l}\Bigr). \tag{2.20}$$

Similarly, the mean value $K_{a,b}(X;\xi,\eta)$ counts the number of integral solutions of the system

$$\sum_{i=1}^k(x_i^j-y_i^j)=\sum_{l=1}^u\sum_{m=1}^k(v_{lm}^j-w_{lm}^j)\quad (1\leqslant j\leqslant k), \tag{2.21}$$

subject to (2.19) and $\mathbf{v}_l,\mathbf{w}_l\in\Xi_b(\eta)\ (\mathrm{mod}\ \varpi^{b+1})$ $(1\leqslant l\leqslant u)$, each solution being counted with weight

$$\rho_a(\xi)^{-2k}\rho_b(\eta)^{-2s}\Bigl(\prod_{i=1}^k\mathfrak{a}_{x_i}\mathfrak{a}_{y_i}\Bigr)\Bigl(\prod_{l=1}^u\prod_{m=1}^k\mathfrak{a}_{v_{lm}}\mathfrak{a}_{w_{lm}}\Bigr). \tag{2.22}$$

Given any one such solution to the system (2.21), an application of the Binomial Theorem shows that $\mathbf{x}-\eta$, $\mathbf{y}-\eta$, $\mathbf{v}-\eta$, $\mathbf{w}-\eta$ is also a solution. Since in any solution counted by $K_{a,b}(X;\xi,\eta)$, one has $\mathbf{v}\equiv\mathbf{w}\equiv\eta\ (\mathrm{mod}\ \varpi^b)$, we deduce in particular that

$$\sum_{i=1}^k(x_i-\eta)^j\equiv\sum_{i=1}^k(y_i-\eta)^j\ (\mathrm{mod}\ \varpi^{jb})\quad (1\leqslant j\leqslant k). \tag{2.23}$$

As in our previous work on efficient congruencing [21], our arguments are considerably simplified by making transparent the relationship between various mean values, on the one hand, and their anticipated magnitudes, on the other. For this reason we consider such mean values normalised by their anticipated orders of magnitude, as follows. We define

$$[[U(X)]]=\frac{U(X)}{X^{s+k-k(k+1)/2}}, \tag{2.24}$$

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

$$[[I_{a,b}(X)]] = \frac{I_{a,b}(X)}{(X/M^a)^{k-k(k+1)/2}(X/M^b)^s}, \tag{2.25}$$

$$[[K_{a,b}(X)]] = \frac{K_{a,b}(X)}{(X/M^a)^{k-k(k+1)/2}(X/M^b)^s}. \tag{2.26}$$

In this notation, our earlier bounds (2.9) and (2.10) for $U(X)$ may be rewritten in the form

$$[[U(X)]] > X^{\Lambda-\delta} \quad \text{and} \quad [[U(Y;\mathfrak{a}')]] < Y^{\Lambda+\delta} \quad (Y \geqslant X^{1/2}), \tag{2.27}$$

in which we write

$$\Lambda = \lambda - (s+k) + k(k+1)/2. \tag{2.28}$$

Our goal is to prove that $\Lambda \leqslant 0$ for $s \geqslant k^2$. This implies that whenever $\varepsilon > 0$ and Z is sufficiently large in terms of ε, s and k, then for any non-zero complex sequence $(\mathfrak{b}_n)$, one has

$$U_{s+k}(Z;\mathfrak{b}) \ll Z^{s+k-k(k+1)/2+\varepsilon}.$$

Thus, whenever $s \geqslant k(k+1)$, $\varepsilon > 0$ and Z is sufficiently large in terms of ε, s and k, it follows that for any complex sequence $(\mathfrak{b}_n)$, one has

$$\oint |f_{\mathfrak{b}}(\boldsymbol{\alpha};Z)|^{2s}\,\mathrm{d}\boldsymbol{\alpha} \ll Z^{s-k(k+1)/2+\varepsilon}\Biggl(\sum_{|n|\leqslant Z}|\mathfrak{b}_n|^2\Biggr)^s,$$

which establishes the first claim of Theorem 1.1. The second claim of Theorem 1.1 follows by application of the Keil-Zhao device, an argument we describe below in §7.

Our strategy may now be outlined in vague terms. We first show that if $U(X)$ is not of the anticipated order of magnitude, then for some $\Lambda > 0$ one has $[[K_{0,1}(X)]] \gg X^{\Lambda-\delta}$, for a suitable large value of X. Next, for a sequence of integers a_n and b_n with b_n roughly equal to ka_n, we show that $[[K_{a_{n+1},b_{n+1}}(X)]]$ is always significantly larger than $[[K_{a_n,b_n}(X)]]$. By iterating this process, we find that for suitably large n, the normalised mean value $[[K_{a_n,b_n}(X)]]$ is so large that we contradict available upper bounds for its value. Thus, one is forced to conclude that $\Lambda \leqslant 0$, as desired.

3. Some consequences of latent translation-dilation invariance

The presence of the coefficients $\mathfrak{a}_n$ prevents direct application of translation-dilation invariance in the mean values $U(X;\mathfrak{a})$. However, our normalisation of the underlying exponential sums $\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)$, and definition via (2.6) of the exponent λ, ensures that much of the power of translation-dilation invariance can nonetheless be extracted. In this section we record the key consequences of this latent translation-dilation invariance for future reference.

We begin with an upper bound for a mean value analogous to $U(X;\mathfrak{a})$.

Lemma 3.1. *Suppose that c is a non-negative integer with $3c\theta \leqslant 1$. Then*

$$\max_{1\leqslant \xi\leqslant \varpi^c} \oint |\mathfrak{f}_c(\boldsymbol{\alpha};\xi)|^{2s+2k}\,\mathrm{d}\boldsymbol{\alpha} \ll (X/M^c)^{\lambda+\delta}. \tag{3.1}$$

Proof. Let ξ be an integer with $1\leqslant \xi\leqslant \varpi^c$. From the definition (2.12) of the exponential sum $\mathfrak{f}_c(\boldsymbol{\alpha};\xi)$, one has

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

in which $\psi(n;\boldsymbol{\alpha})$ is given by (2.3), and

$$\mathfrak{b}_y(\xi)=\mathfrak{a}_{\varpi^c y+\xi}. \tag{3.2}$$

By orthogonality, one finds that the integral on the left hand side of (3.1) counts the number of integral solutions of the system of equations

$$\sum_{i=1}^{s+k}(\varpi^c y_i+\xi)^j=\sum_{i=1}^{s+k}(\varpi^c z_i+\xi)^j\quad (1\leqslant j\leqslant k), \tag{3.3}$$

with $-(X+\xi)/\varpi^c\leqslant \mathbf{y},\mathbf{z}\leqslant (X-\xi)/\varpi^c$, each solution being counted with weight

$$\rho_c(\xi)^{-2s-2k}\mathfrak{b}_{y_1}\dots \mathfrak{b}_{y_{s+k}}\mathfrak{b}_{z_1}\dots \mathfrak{b}_{z_{s+k}}.$$

An application of the Binomial Theorem shows that the pair $\mathbf{y},\mathbf{z}$ satisfies (3.3) if and only if it satisfies the system

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

Thus, recalling (2.2), reversing track and accommodating the end-points of the summation, we find that

$$\oint |\mathfrak{f}_c(\boldsymbol{\alpha};\xi)|^{2s+2k}\,\mathrm{d}\boldsymbol{\alpha}=\oint |\widetilde{f}_{\mathfrak{b}'}(\boldsymbol{\alpha};(X+\xi)/\varpi^c)|^{2s+2k}\,\mathrm{d}\boldsymbol{\alpha}, \tag{3.4}$$

where

$$\mathfrak{b}'_y=\begin{cases}\mathfrak{b}_y(\xi), & \text{when } -(X+\xi)/\varpi^c\leqslant y\leqslant (X-\xi)/\varpi^c,\\ 0, & \text{when } (X-\xi)/\varpi^c< y\leqslant (X+\xi)/\varpi^c.\end{cases}$$

Here, we have made the trivial observation that, in view of the relation (3.2), one has

$$\begin{aligned}\rho(\mathfrak{b}';(X+\xi)/\varpi^c)^2&=\sum_{-(X+\xi)/\varpi^c\leqslant y\leqslant (X-\xi)/\varpi^c}|\mathfrak{b}_y(\xi)|^2\\ &=\sum_{\substack{|n|\leqslant X\\ n\equiv \xi \ (\mathrm{mod}\ \varpi^c)}}|\mathfrak{a}_n|^2=\rho_c(\xi)^2.\end{aligned}$$

Observe next that the hypothesis $3c\theta \leqslant 1$ ensures that $(X+\xi)/\varpi^c \geqslant X/M^c > X^{1/2}$. Thus, the upper bound (2.10) supplies the estimate

$$\oint |\widetilde{f}_{\mathfrak{b}'}(\boldsymbol{\alpha};(X+\xi)/\varpi^c)|^{2s+2k}\,\mathrm{d}\boldsymbol{\alpha} = U((X+\xi)/\varpi^c;\mathfrak{b}')$$
$$< ((X+\xi)/\varpi^c)^{\lambda+\delta} \ll (X/M^c)^{\lambda+\delta}.$$

The desired conclusion now follows by substituting this bound into (3.4). □

We record an additional estimate to demystify a bound of which we make use in our discussion of the congruencing step.

Lemma 3.2. *Suppose that c is a non-negative integer with $3c\theta \leqslant 1$. Then*

$$\max_{1\leqslant \xi\leqslant \varpi^c} \oint |\mathfrak{F}_c(\boldsymbol{\alpha};\xi)|^{2u+2}\,\mathrm{d}\boldsymbol{\alpha} \ll (X/M^c)^{\lambda+\delta}.$$

Proof. From the definitions (2.12) and (2.13) of $\mathfrak{f}_c(\boldsymbol{\alpha};\xi)$ and $\mathfrak{F}_c(\boldsymbol{\alpha};\xi)$, one has

$$\oint |\mathfrak{F}_c(\boldsymbol{\alpha};\xi)|^{2u+2}\,\mathrm{d}\boldsymbol{\alpha} = \rho_c(\xi)^{-2s-2k}\oint \Bigl| \sum_{\boldsymbol{\xi}\in\Xi_c(\xi)} \prod_{i=1}^{k} \mathfrak{g}_{c+1}(\boldsymbol{\alpha};\xi_i)\Bigr|^{2u+2}\mathrm{d}\boldsymbol{\alpha}, \qquad (3.5)$$

where we temporarily write

$$\mathfrak{g}_d(\boldsymbol{\alpha};\zeta) = \sum_{\substack{|n|\leqslant X\\ n\equiv \zeta \ (\mathrm{mod}\ \varpi^d)}} \mathfrak{a}_n e(\psi(n;\boldsymbol{\alpha})).$$

We presume that the weights $\mathfrak{a}_n$ are positive, and hence the mean value on the right hand side of (3.5) counts the number of integral solutions of the system of equations

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

with $|\mathbf{x}|,|\mathbf{y}| \leqslant X$, $\mathbf{x}\equiv\mathbf{y}\equiv\xi \pmod{\varpi^c}$ and certain additional congruence conditions imposed by conditioning hypotheses, each solution being counted with the non-negative weight

$$\rho_c(\xi)^{-2s-2k}\mathfrak{a}_{x_1}\dots\mathfrak{a}_{x_{s+k}}\mathfrak{a}_{y_1}\dots\mathfrak{a}_{y_{s+k}}.$$

Since these weights are non-negative, the omission of the additional congruence conditions cannot decrease the resulting estimate, and thus

$$\oint |\mathfrak{F}_c(\boldsymbol{\alpha};\xi)|^{2u+2}\,\mathrm{d}\boldsymbol{\alpha} \leqslant \rho_c(\xi)^{-2s-2k}\oint |\mathfrak{g}_c(\boldsymbol{\alpha};\xi)|^{2s+2k}\,\mathrm{d}\boldsymbol{\alpha}$$
$$= \oint |\mathfrak{f}_c(\boldsymbol{\alpha};\xi)|^{2s+2k}\,\mathrm{d}\boldsymbol{\alpha}.$$

The conclusion of the lemma is now immediate from Lemma 3.1. □

A variant of Lemma 3.1 proves useful both in this section and elsewhere.

Lemma 3.3. *Suppose that c and d are non-negative integers satisfying the condition* $\max\{3c\theta, 3d\theta\} \leqslant 1$. *Then*

$$\max_{1\leqslant \xi\leqslant \varpi^c}\max_{1\leqslant \eta\leqslant \varpi^d}\oint |\mathfrak{f}_c(\boldsymbol{\alpha};\xi)^{2k}\mathfrak{f}_d(\boldsymbol{\alpha};\eta)^{2s}|\,\mathrm{d}\boldsymbol{\alpha} \ll \bigl((X/M^c)^k(X/M^d)^s\bigr)^{(\lambda+\delta)/(s+k)}.$$

Proof. Let ξ and η be integers with $1\leqslant \xi\leqslant \varpi^c$ and $1\leqslant \eta\leqslant \varpi^d$. An application of Hölder's inequality reveals that

$$\oint |\mathfrak{f}_c(\boldsymbol{\alpha};\xi)^{2k}\mathfrak{f}_d(\boldsymbol{\alpha};\eta)^{2s}|\,\mathrm{d}\boldsymbol{\alpha}$$
$$\leqslant \Bigl(\oint |\mathfrak{f}_c(\boldsymbol{\alpha};\xi)|^{2s+2k}\,\mathrm{d}\boldsymbol{\alpha}\Bigr)^{k/(s+k)}\Bigl(\oint |\mathfrak{f}_d(\boldsymbol{\alpha};\eta)|^{2s+2k}\,\mathrm{d}\boldsymbol{\alpha}\Bigr)^{s/(s+k)}.$$

The conclusion of the lemma follows from this bound by inserting the estimate supplied by Lemma 3.1 to estimate the mean values occurring on the right hand side. □

The last lemma of this section provides a crude estimate for the quantity $K_{a,b}(X)$ of use at the end of our iterative process.

Lemma 3.4. *Suppose that a and b are integers with $0\leqslant a<b\leqslant (3\theta)^{-1}$. Then provided that $\Lambda\geqslant 0$, one has*

$$[\![K_{a,b}(X)]\!] \ll X^{\Lambda+\delta}(M^{b-a})^{k(k+1)/2}.$$

Proof. Let ξ and η be integers with $1\leqslant \xi\leqslant \varpi^c$ and $1\leqslant \eta\leqslant \varpi^d$. Then, as in the discussion of §2, we find that the mean value $K_{a,b}(X;\xi,\eta)$ counts the number of integral solutions $\mathbf{x},\mathbf{y},\mathbf{v},\mathbf{w}$ of the system (2.21) subject to $|\mathbf{x}|,|\mathbf{y}|,|\mathbf{v}|,|\mathbf{w}|\leqslant X$,

$$\mathbf{x},\mathbf{y}\in \Xi_a(\xi)\ (\mathrm{mod}\ \varpi^{a+1}),\quad \mathbf{v}_l,\mathbf{w}_l\in \Xi_b(\eta)\ (\mathrm{mod}\ \varpi^{b+1})\quad (1\leqslant l\leqslant u),\qquad (3.6)$$

each solution being counted with weight (2.22). Since we suppose the weights $\mathfrak{a}_n$ to be positive, we may relax the conditions (3.6) to insist only that

$$\mathbf{x}\equiv \mathbf{y}\equiv \xi\ (\mathrm{mod}\ \varpi^a),\quad \mathbf{v}_l\equiv \mathbf{w}_l\equiv \eta\ (\mathrm{mod}\ \varpi^b)\quad (1\leqslant l\leqslant u),$$

this relaxation only increasing our resulting estimate for $K_{a,b}(X;\xi,\eta)$. By reinterpreting the associated number of solutions of the system (2.21) via orthogonality and invoking Lemma 3.3, we deduce that

$$K_{a,b}(X;\xi,\eta)\leqslant \oint |\mathfrak{f}_a(\boldsymbol{\alpha};\xi)^{2k}\mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{2s}|\,\mathrm{d}\boldsymbol{\alpha}$$
$$\ll \bigl((X/M^a)^k(X/M^b)^s\bigr)^{(\lambda+\delta)/(s+k)}.$$

Next we recall (2.17), and note that the definition (2.11) ensures that

$$\sum_{\xi=1}^{\varpi^a}\rho_a(\xi)^2=\rho_0(1)^2=\sum_{\eta=1}^{\varpi^b}\rho_b(\eta)^2.$$

In view of (2.28), we deduce that

$$K_{a,b}(X)\ll (X/M^a)^{k-k(k+1)/2}(X/M^b)^s\bigl(M^{s(b-a)/(s+k)}\bigr)^{k(k+1)/2}X^{\Lambda+\delta}.$$

We consequently conclude from (2.26) that

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

which suffices to complete the proof of the lemma. $\square$

4. The conditioning process

The variables underlying the mean value $I_{a,b}(X;\xi,\eta)$ must be conditioned so as to ensure that appropriate non-singularity conditions hold, yielding the mean value $K_{a,b}(X;\xi,\eta)$ central to the congruencing process. This we achieve in the next two lemmata. We note in this context that our conditioning treatment here is considerably sharper than in our earlier work associated with Vinogradov's mean value theorem.

Lemma 4.1. *Let a and b be integers with $b > a \geqslant 0$. Then one has*

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

Proof. Consider fixed integers ξ and η with $1 \leqslant \xi \leqslant \varpi^a$ and $1 \leqslant \eta \leqslant \varpi^b$. Then by orthogonality, one finds from (2.14) that $I_{a,b}(X;\xi,\eta)$ counts the number of integral solutions $\mathbf{x}, \mathbf{y}, \mathbf{v}, \mathbf{w}$ of the system (2.18), with its attendant conditions, and with each solution counted with weight (2.20). Let $T_1(\xi,\eta)$ denote the contribution to $I_{a,b}(X;\xi,\eta)$ arising from those integral solutions in which three at least of the integers $v_1, \ldots, v_s$ lie in a common congruence class modulo ϖ^{b+1}. Also, let $T_2(\xi,\eta)$ denote the corresponding contribution to $I_{a,b}(X;\xi,\eta)$ arising from those integral solutions in which no three of the integers $v_1, \ldots, v_s$ lie in a common congruence class modulo ϖ^{b+1}. Thus we have

$$I_{a,b}(X;\xi,\eta) = T_1(\xi,\eta) + T_2(\xi,\eta). \tag{4.1}$$

We begin by estimating the quantity $T_1(\xi,\eta)$. Let $\mathbf{x}, \mathbf{y}, \mathbf{v}, \mathbf{w}$ be a solution contributing to $T_1(\xi,\eta)$. By relabelling the suffices of the variables $v_1, \ldots, v_s$, we may suppose that $v_1 \equiv v_2 \equiv v_3 \pmod{\varpi^{b+1}}$, provided that we inflate the ensuing estimates by a factor $\binom{s}{3}$. Then, on recalling the presumed positivity of the weights $\mathfrak{a}_n$, it follows via orthogonality that

$$T_1(\xi,\eta) \ll \oint |\mathfrak{F}_a(\boldsymbol{\alpha};\xi)^2 \mathfrak{G}_b(\boldsymbol{\alpha};\eta) \mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{2s-3}| \,\mathrm{d}\boldsymbol{\alpha}, \tag{4.2}$$

in which we write

$$\mathfrak{G}_b(\boldsymbol{\alpha};\eta) = \rho_b(\eta)^{-3} \Biggl(\sum_{\substack{1 \leqslant \zeta \leqslant \varpi^{b+1} \\ \zeta \equiv \eta \pmod{\varpi^b}}} \rho_{b+1}(\zeta) \mathfrak{f}_{b+1}(\boldsymbol{\alpha};\zeta) \Biggr)^3. \tag{4.3}$$

We note in this context that our assumption $s \geqslant k^2$ ensures that $s \geqslant 3$. In view of (2.14), an application of Hölder's inequality on the right hand side of (4.2) yields the bound

$$T_1(\xi,\eta) \ll I_{a,b}(X;\xi,\eta)^{1-3/(2s)} \biggl(\oint |\mathfrak{F}_a(\boldsymbol{\alpha};\xi)^2 \mathfrak{G}_b(\boldsymbol{\alpha};\eta)^{2s/3}| \,\mathrm{d}\boldsymbol{\alpha} \biggr)^{3/(2s)}.$$

Thus, now referring to (2.16), a second application of Hölder's inequality leads from (4.1) to the estimate

$$I_{a,b}(X) \ll I_{a,b}(X)^{1-3/(2s)} T_3^{3/(2s)} + \rho_0(1)^{-4} \sum_{\xi=1}^{\varpi^a} \sum_{\eta=1}^{\varpi^b} \rho_a(\xi)^2 \rho_b(\eta)^2 T_2(\xi,\eta),$$

where

$$T_3 = \rho_0(1)^{-4} \sum_{\xi=1}^{\varpi^a} \sum_{\eta=1}^{\varpi^b} \rho_a(\xi)^2 \rho_b(\eta)^2 \oint |\mathfrak{F}_a(\boldsymbol{\alpha};\xi)^2 \mathfrak{G}_b(\boldsymbol{\alpha};\eta)^{2s/3}| \,\mathrm{d}\boldsymbol{\alpha}. \tag{4.4}$$

We therefore deduce that

$$I_{a,b}(X) \ll T_3 + \rho_0(1)^{-4} \sum_{\xi=1}^{\varpi^a} \sum_{\eta=1}^{\varpi^b} \rho_a(\xi)^2 \rho_b(\eta)^2 T_2(\xi,\eta). \tag{4.5}$$

In order to estimate T_3, we begin with an application of Hölder's inequality to (4.3), obtaining

$$|\mathfrak{G}_b(\boldsymbol{\alpha};\eta)|^{2s/3} \leqslant \left(\mathfrak{H}_b^{(1)}(\boldsymbol{\alpha};\eta)\right)^{s(2s-3)/(3s-3)} \left(\mathfrak{H}_b^{(s)}(\boldsymbol{\alpha};\eta)\right)^{s/(3s-3)},$$

where we write

$$\mathfrak{H}_b^{(t)}(\boldsymbol{\alpha};\eta) = \rho_b(\eta)^{-2t} \sum_{\substack{1\leqslant \zeta \leqslant \varpi^{b+1} \\ \zeta\equiv \eta \ (\mathrm{mod}\ \varpi^b)}} \rho_{b+1}(\zeta)^{2t} |\mathfrak{f}_{b+1}(\boldsymbol{\alpha};\zeta)|^{2t}. \tag{4.6}$$

A further application of Hölder's inequality consequently conveys us from (4.4) to the bound

$$T_3 \leqslant T_4^{(2s-3)/(3s-3)} T_5^{s/(3s-3)}, \tag{4.7}$$

in which

$$T_4 = \rho_0(1)^{-4} \sum_{\xi=1}^{\varpi^a} \sum_{\eta=1}^{\varpi^b} \rho_a(\xi)^2 \rho_b(\eta)^2 \oint |\mathfrak{F}_a(\boldsymbol{\alpha};\xi)^2 \mathfrak{H}_b^{(1)}(\boldsymbol{\alpha};\eta)^s| \,\mathrm{d}\boldsymbol{\alpha} \tag{4.8}$$

and

$$T_5 = \rho_0(1)^{-4} \sum_{\xi=1}^{\varpi^a} \sum_{\eta=1}^{\varpi^b} \rho_a(\xi)^2 \rho_b(\eta)^2 \oint |\mathfrak{F}_a(\boldsymbol{\alpha};\xi)^2 \mathfrak{H}_b^{(s)}(\boldsymbol{\alpha};\eta)| \,\mathrm{d}\boldsymbol{\alpha}. \tag{4.9}$$

The integral within the definition (4.8) of T_4 counts the number of integral solutions of the system (2.18), subject to its attendant conditions, with weight (2.20), and subject to the additional condition $v_l \equiv w_l \pmod{\varpi^{b+1}}$ $(1 \leqslant l \leqslant s)$. Since the weights $\mathfrak{a}_n$ are presumed positive, the omission of this last condition merely inflates our estimate for this integral, and thus we see that it is bounded above by $I_{a,b}(X;\xi,\eta)$. We thus deduce that

$$T_4 \leqslant I_{a,b}(X). \tag{4.10}$$

In order to estimate T_5, we again recall that the weights $\mathfrak{a}_n$ are presumed positive. Thus, when $\zeta \equiv \eta \pmod{\varpi^b}$, we have the upper bound

$$\rho_{b+1}(\zeta)^2 = \sum_{\substack{|n|\leqslant X\\ n\equiv \zeta \ (\mathrm{mod}\ \varpi^{b+1})}} |\mathfrak{a}_n|^2 \leqslant \sum_{\substack{|n|\leqslant X\\ n\equiv \eta \ (\mathrm{mod}\ \varpi^{b})}} |\mathfrak{a}_n|^2 = \rho_b(\eta)^2,$$

so that (4.6) yields the bound

$$\mathfrak{H}_b^{(s)}(\boldsymbol{\alpha};\eta) \leqslant \rho_b(\eta)^{-2} \sum_{\substack{1\leqslant \zeta\leqslant \varpi^{b+1}\\ \zeta\equiv \eta \ (\mathrm{mod}\ \varpi^{b})}} \rho_{b+1}(\zeta)^2 |\mathfrak{f}_{b+1}(\boldsymbol{\alpha};\zeta)|^{2s}.$$

On substituting this bound into (4.9), we infer from (2.14) that

$$\begin{aligned} T_5 &\leqslant \rho_0(1)^{-4} \sum_{\xi=1}^{\varpi^a} \sum_{\eta=1}^{\varpi^b} \sum_{\substack{1\leqslant \zeta\leqslant \varpi^{b+1}\\ \zeta\equiv \eta \ (\mathrm{mod}\ \varpi^{b})}} \rho_a(\xi)^2 \rho_{b+1}(\zeta)^2 I_{a,b+1}(X;\xi,\zeta)\\ &= \rho_0(1)^{-4} \sum_{\xi=1}^{\varpi^a} \sum_{\zeta=1}^{\varpi^{b+1}} \rho_a(\xi)^2 \rho_{b+1}(\zeta)^2 I_{a,b+1}(X;\xi,\zeta). \end{aligned}$$

On recalling (2.16), we therefore see that $T_5 \leqslant I_{a,b+1}(X)$, and thus we deduce from (4.7) and (4.10) that

$$T_3 \leqslant (I_{a,b}(X))^{(2s-3)/(3s-3)} (I_{a,b+1}(X))^{s/(3s-3)}.$$

Substituting this bound into (4.5) and disentangling the result, we conclude thus far that

$$I_{a,b}(X) \ll I_{a,b+1}(X) + \rho_0(1)^{-4} \sum_{\xi=1}^{\varpi^a} \sum_{\eta=1}^{\varpi^b} \rho_a(\xi)^2 \rho_b(\eta)^2 T_2(\xi,\eta). \tag{4.11}$$

Now is the moment to estimate the contribution of $T_2(\xi,\eta)$. Let $\mathbf{x},\mathbf{y},\mathbf{v},\mathbf{w}$ be a solution contributing to $T_2(\xi,\eta)$. Since $s \geqslant k^2 \geqslant 2k$, and no three of the integers $v_1,\ldots,v_s$ lie in a common congruence class modulo ϖ^{b+1}, it follows that $v_1,\ldots,v_s$ together occupy at least k distinct congruence classes modulo ϖ^{b+1}. By relabelling the suffices of the variables $v_1,\ldots,v_s$, we may suppose that $v_1,\ldots,v_k$ lie in distinct congruence classes modulo ϖ^{b+1}, provided that we inflate the ensuing estimates by a factor $\binom{s}{k}$. Then, again recalling that the weights $\mathfrak{a}_n$ may be assumed positive, it follows via orthogonality that

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

By reference to (2.14) and (2.15), an application of Hölder's inequality shows that

$$T_2(\xi,\eta) \ll I_{a,b}(X;\xi,\eta)^{1-k/(2s)} K_{a,b}(X;\xi,\eta)^{k/(2s)}.$$

Here we recall that $s=uk$. Thus, appealing to Hölder's inequality yet again, we conclude via (2.16) and (2.17) that

$$\rho_0(1)^{-4}\sum_{\xi=1}^{\varpi^a}\sum_{\eta=1}^{\varpi^b}\rho_a(\xi)^2\rho_b(\eta)^2T_2(\xi,\eta)\ll I_{a,b}(X)^{1-k/(2s)}K_{a,b}(X)^{k/(2s)}.$$

On substituting this bound into (4.11), we obtain

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

and the conclusion of the lemma follows by disentangling. □

By iterating Lemma 4.1, we are able to estimate $I_{a,b}(X)$ in terms of conditioned mean values of type $K_{a,b+h}(X)$ $(h\geqslant 0)$.

Lemma 4.2. *Let a and b be integers with $0\leqslant a<b$, and put $H=4(b-a)$. Suppose that $b+H\leqslant (3\theta)^{-1}$. Then there exists an integer h with $0\leqslant h<H$ having the property that*

$$I_{a,b}(X)\ll K_{a,b+h}(X)+M^{-sH/4}(X/M^b)^s(X/M^a)^{k-k(k+1)/2+\Lambda}.$$

Proof. By repeated application of Lemma 4.1, we obtain the bound

$$I_{a,b}(X)\ll \sum_{h=0}^{H-1}K_{a,b+h}(X)+I_{a,b+H}(X). \tag{4.12}$$

Let ξ and η be fixed integers with $1\leqslant \xi\leqslant \varpi^a$ and $1\leqslant \eta\leqslant \varpi^{b+H}$. Then on recalling our presumption that the weights $\mathfrak{a}_n$ are positive, it follows from (2.14) via orthogonality that

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

We therefore deduce from Lemma 3.3 that

$$\begin{aligned} I_{a,b+H}(X;\xi,\eta)&\ll \left((X/M^a)^k(X/M^{b+H})^s\right)^{(\lambda+\delta)/(s+k)}\\ &\ll X^\delta(X/M^a)^{\lambda-s}(X/M^b)^sM^\Omega, \end{aligned}\tag{4.13}$$

in which

$$\Omega=\lambda\left(a-\frac{ak}{s+k}-\frac{bs}{s+k}\right)+s(b-a)-\frac{Hs\lambda}{s+k}.$$

We may suppose that $\lambda\geqslant s+k-k(k+1)/2\geqslant \tfrac12(s+k)$. Then since $b\geqslant a$, we obtain the estimate

$$\Omega\leqslant -\frac{s(b-a)\lambda}{s+k}+s(b-a)-\tfrac12 Hs\leqslant \tfrac12 s(b-a-H).$$

But $H=4(b-a)$, and so we discern from (2.8) that

$$\Omega\leqslant -\tfrac38 Hs\leqslant -\delta\theta^{-1}-\tfrac14 Hs.$$

Substituting this estimate into (4.13), we see that

$$I_{a,b+H}(X;\xi,\eta)\ll M^{-sH/4}(X/M^a)^{\lambda-s}(X/M^b)^s.$$

We next recall (2.16), deducing that

$$I_{a,b+H}(X) \ll \rho M^{-sH/4}(X/M^a)^{\lambda-s}(X/M^b)^s,$$

where

$$\rho = \rho_0(1)^{-4}\left(\sum_{\xi=1}^{\varpi^a}\rho_a(\xi)^2\right)\left(\sum_{\eta=1}^{\varpi^{b+H}}\rho_{b+H}(\eta)^2\right) = 1.$$

The conclusion of the lemma consequently follows from (4.12). $\square$

We next introduce a lemma that initiates the iterative process.

Lemma 4.3. *There exists a prime ϖ with $M < \varpi \leqslant 2M$ for which*

$$U(X) \ll M^s I_{0,1}(X).$$

Proof. By orthogonality, it follows from (2.4) that $U(X)$ counts the number of integral solutions $\mathbf{x}, \mathbf{y}$ of the system

$$\sum_{i=1}^{s+k}(x_i^j - y_i^j) = 0 \quad (1 \leqslant j \leqslant k), \tag{4.14}$$

with $|\mathbf{x}|, |\mathbf{y}| \leqslant X$, each solution being counted with weight

$$\rho_0(1)^{-2s-2k}\prod_{i=1}^{s+k}\mathfrak{a}_{x_i}\mathfrak{a}_{y_i}.$$

Let T_0 denote the contribution 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 contribution with $x_i = x_j$ for no i and j with $1 \leqslant i < j \leqslant k$. Then we have

$$U(X) = T_0 + T_1. \tag{4.15}$$

We recall again that the weights $\mathfrak{a}_n$ are presumed positive. Then by relabelling the suffices of $x_1, \ldots, x_k$, we find that T_0 is bounded above by $\binom{k}{2}T_2$, where T_2 denotes the number of solutions of the system

$$2x_1^j + \sum_{i=3}^{s+k}x_i^j = \sum_{l=1}^{s+k}y_l^j \quad (1 \leqslant j \leqslant k),$$

with $|\mathbf{x}|, |\mathbf{y}| \leqslant X$, each solution being counted with weight

$$\rho_0(1)^{-2s-2k}\mathfrak{a}_{x_1}^2\left(\prod_{i=3}^{s+k}\mathfrak{a}_{x_i}\right)\left(\prod_{l=1}^{s+k}\mathfrak{a}_{y_l}\right).$$

Put $\mathfrak{b}_n = \mathfrak{a}_n^2$ for $|n| \leqslant X$. Then on recalling (1.2), it follows via orthogonality and the triangle inequality that

$$T_0 \ll \rho_0(1)^{-2}\oint |f_{\mathfrak{b}}(2\boldsymbol{\alpha}; X)\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha}; X)^{2s+2k-2}|\,\mathrm{d}\boldsymbol{\alpha}.$$

A second application of the triangle inequality reveals that

$$|f_{\mathfrak{b}}(2\boldsymbol{\alpha}; X)| \leqslant \sum_{|n|\leqslant X}|\mathfrak{b}_n| = \sum_{|n|\leqslant X}|\mathfrak{a}_n|^2 = \rho_0(1)^2.$$

Thus, an application of Hölder's inequality gives the bound

$$T_0 \ll \left(\oint |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha}; X)|^{2s+2k} \, \mathrm{d}\boldsymbol{\alpha} \right)^{1-1/(s+k)} = U(X)^{1-1/(s+k)}.$$

On substituting this estimate into (4.15) and disentangling, we deduce that

$$U(X) \ll 1 + T_1. \tag{4.16}$$

Consider next a solution $\mathbf{x}, \mathbf{y}$ of (4.14) contributing to 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 the set consisting of the smallest $[k^3/\theta] + 1$ prime numbers exceeding M. It follows from the Prime Number Theorem that none of these primes exceed $2M$. Moreover, one has

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

and hence one at least of the primes belonging to $\mathcal{P}$ does not divide $\Delta(\mathbf{x})$. In particular, there is a prime $\varpi \in \mathcal{P}$ for which $x_i \equiv x_j \pmod{\varpi}$ for no i and j with $1 \leqslant i < j \leqslant k$. Again recalling that the weights $\mathfrak{a}_n$ are presumed positive, it follows from (2.13) via orthogonality that

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

Therefore, as a consequence of Schwarz's inequality, one finds via (2.4) and (2.14) that there exists a prime $\varpi \in \mathcal{P}$ for which

$$\begin{aligned} T_1 &\ll \left(\oint |\mathfrak{F}_0(\boldsymbol{\alpha}; 1)^2 \mathfrak{f}_0(\boldsymbol{\alpha}; 1)^{2s}| \, \mathrm{d}\boldsymbol{\alpha} \right)^{1/2} \left(\oint |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha}; X)|^{2s+2k} \, \mathrm{d}\boldsymbol{\alpha} \right)^{1/2} \\ &= I_{0,0}(X; 1, 1)^{1/2} U(X)^{1/2}. \end{aligned}$$

Thus we conclude by means of (4.16) that

$$U(X) \ll 1 + I_{0,0}(X; 1, 1)^{1/2} U(X)^{1/2},$$

whence

$$U(X) \ll I_{0,0}(X; 1, 1). \tag{4.17}$$

Here, we have made use of a trivial lower bound for $I_{0,0}(X; 1, 1)$ obtained by considering the diagonal contribution in combination with the assumption that the coefficients $\mathfrak{a}_n$ are positive.

Next, we split the summation in (2.2) into arithmetic progressions modulo ϖ. Thus we obtain

$$\mathfrak{f}_0(\boldsymbol{\alpha}; 1) = \rho_0(1)^{-1} \sum_{\xi=1}^{\varpi} \rho_1(\xi) \mathfrak{f}_1(\boldsymbol{\alpha}; \xi).$$

By applying Hölder's inequality, we deduce that

$$|\mathfrak{f}_0(\boldsymbol{\alpha};1)|^{2s} \leqslant \rho_0(1)^{-2s}\biggl(\sum_{\xi=1}^{\varpi}\rho_1(\xi)^2\biggr)^{s-1}\biggl(\sum_{\xi=1}^{\varpi}1\biggr)^{s}\biggl(\sum_{\xi=1}^{\varpi}\rho_1(\xi)^2|\mathfrak{f}_1(\boldsymbol{\alpha};\xi)|^{2s}\biggr).$$

But

$$\sum_{\xi=1}^{\varpi}\rho_1(\xi)^2=\rho_0(1)^2,$$

and hence we deduce from (2.14) and (2.16) that

$$\begin{aligned} I_{0,0}(X;1,1) &\ll \rho_0(1)^{-2}M^s\sum_{\xi=1}^{\varpi}\rho_1(\xi)^2\oint|\mathfrak{F}_0(\boldsymbol{\alpha};1)^2\mathfrak{f}_1(\boldsymbol{\alpha};\xi)^{2s}|\,\mathrm{d}\boldsymbol{\alpha}\\ &=\rho_0(1)^{-2}M^s\sum_{\xi=1}^{\varpi}\rho_1(\xi)^2I_{0,1}(X;\xi,1)\\ &=M^sI_{0,1}(X). \end{aligned}$$

The conclusion of the lemma is now immediate from (4.17). □

We now fix the prime ϖ, once and for all, in accordance with Lemma 4.3. Finally, we obtain the starting point of our iteration in the next lemma.

Lemma 4.4. *There exists an integer $h\in\{0,1,2,3\}$ for which one has*

$$U(X)\ll M^sK_{0,1+h}(X).$$

Proof. According to Lemma 4.2, there exists an integer h with $0\leqslant h<4$ having the property that

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

Since we may suppose that $M>X^\delta$, it follows from Lemma 4.3 that

$$U(X)\ll M^sK_{0,1+h}(X)+X^{s+k-k(k+1)/2+\Lambda-2\delta}.$$

But in view of (2.27), we have

$$U(X)>X^{s+k-k(k+1)/2+\Lambda-\delta},$$

and hence we arrive at the upper bound

$$U(X)\ll M^sK_{0,1+h}(X)+X^{-\delta}U(X).$$

The conclusion of the lemma is now immediate. □

5. The efficient congruencing step

The mean value $K_{a,b}(X;\xi,\eta)$ is subject to a powerful congruence condition on its underlying variables. In this section, we uncover this condition and convert it into one suitable for iteration. We begin with some preliminary discussion of congruences. Denote by $\mathcal{B}_{a,b}(\mathbf{m};\xi,\eta)$ the set of solutions of the system of congruences

$$\sum_{i=1}^k (z_i-\eta)^j \equiv m_j \pmod{\varpi^{jb}} \quad (1\leqslant j\leqslant k), \tag{5.1}$$

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

Lemma 5.1. *Suppose that a and b are non-negative integers with $b>a$. Then*

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

Proof. This is a special case of [21, Lemma 4.1]. □

Lemma 5.2. *Suppose that a and b are integers with $0\leqslant a<b\leqslant \theta^{-1}$. Then*

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

Proof. Let ξ and η be fixed integers with $1\leqslant \xi\leqslant \varpi^a$ and $1\leqslant \eta\leqslant \varpi^b$. We recall that $K_{a,b}(X;\xi,\eta)$ counts the number of integral solutions $\mathbf{x},\mathbf{y},\mathbf{v},\mathbf{w}$ of the system (2.21), subject to its attendant conditions, with each solution counted with weight (2.22). Moreover, these solutions are subject to the congruence conditions (2.23). A comparison between the latter and (5.1) shows that for each solution $\mathbf{x},\mathbf{y}$ of the system of congruences (2.23), there is a k-tuple $\mathbf{m}$ for which one has both $\mathbf{x}\in \mathcal{B}_{a,b}(\mathbf{m};\xi,\eta)$ and $\mathbf{y}\in \mathcal{B}_{a,b}(\mathbf{m};\xi,\eta)$.

Write

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

Then it follows by orthogonality that

$$K_{a,b}(X;\xi,\eta)=\sum_{m_1=1}^{\varpi^b}\dots\sum_{m_k=1}^{\varpi^{kb}}\oint |\mathfrak{G}_{a,b}(\boldsymbol{\alpha};\xi,\eta;\mathbf{m})^2\mathfrak{F}_b(\boldsymbol{\alpha};\eta)^{2u}|\,\mathrm{d}\boldsymbol{\alpha}. \tag{5.2}$$

An application of Cauchy's inequality in combination with Lemma 5.1 yields the bound

$$\begin{aligned}|\mathfrak{G}_{a,b}(\boldsymbol{\alpha};\xi,\eta;\mathbf{m})|^2&\leqslant \mathrm{card}(\mathcal{B}_{a,b}(\mathbf{m};\xi,\eta))\sum_{\boldsymbol{\zeta}\in \mathcal{B}_{a,b}(\mathbf{m};\xi,\eta)}\prod_{i=1}^k \frac{\rho_{kb}(\zeta_i)^2}{\rho_a(\xi)^2}|\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta_i)|^2\\ &\ll M^{\frac{1}{2}k(k-1)(a+b)}\sum_{\boldsymbol{\zeta}\in \mathcal{B}_{a,b}(\mathbf{m};\xi,\eta)}\prod_{i=1}^k \frac{\rho_{kb}(\zeta_i)^2}{\rho_a(\xi)^2}|\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta_i)|^2. \end{aligned}\tag{5.3}$$

On substituting (5.3) into (5.2), and again applying orthogonality, we deduce that

$$K_{a,b}(X;\xi,\eta)\ll M^{\frac{1}{2}k(k-1)(a+b)}\sum_{\substack{1\leqslant \boldsymbol{\zeta}\leqslant \varpi^{kb}\\ \boldsymbol{\zeta}\equiv \xi \pmod{\varpi^a}}}\oint |\mathfrak{F}_b(\boldsymbol{\alpha};\eta)|^{2u}\prod_{i=1}^k \frac{\rho_{kb}(\zeta_i)^2}{\rho_a(\xi)^2}|\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta_i)|^2. \tag{5.4}$$

Here, we have again made use of the presumed positivity of the coefficients $\mathfrak{a}_n$ in order to drop the implicit conditioning of the final block of $2k$ variables.

Next we observe that an application of Hölder's inequality reveals that

$$\sum_{\substack{1\leqslant \boldsymbol{\zeta}\leqslant \varpi^{kb}\\ \boldsymbol{\zeta}\equiv \xi \ (\mathrm{mod}\ \varpi^a)}} \prod_{i=1}^k \frac{\rho_{kb}(\zeta_i)^2}{\rho_a(\xi)^2}|\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta_i)|^2 = \Biggl(\sum_{\substack{1\leqslant \zeta\leqslant \varpi^{kb}\\ \zeta\equiv \xi \ (\mathrm{mod}\ \varpi^a)}} \frac{\rho_{kb}(\zeta)^2}{\rho_a(\xi)^2}|\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta)|^2\Biggr)^k$$
$$\leqslant \rho^{k-1} \sum_{\substack{1\leqslant \zeta\leqslant \varpi^{kb}\\ \zeta\equiv \xi \ (\mathrm{mod}\ \varpi^a)}} \frac{\rho_{kb}(\zeta)^2}{\rho_a(\xi)^2}|\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta)|^{2k},$$

where

$$\rho = \sum_{\substack{1\leqslant \zeta\leqslant \varpi^{kb}\\ \zeta\equiv \xi \ (\mathrm{mod}\ \varpi^a)}} \frac{\rho_{kb}(\zeta)^2}{\rho_a(\xi)^2} = \rho_a(\xi)^{-2} \sum_{\substack{|n|\leqslant X\\ n\equiv \xi \ (\mathrm{mod}\ \varpi^a)}} |\mathfrak{a}_n|^2 = 1.$$

Thus we deduce from (5.4) that

$$K_{a,b}(X;\xi,\eta) \ll M^{\frac{1}{2}k(k-1)(a+b)} \sum_{\substack{1\leqslant \zeta\leqslant \varpi^{kb}\\ \zeta\equiv \xi \ (\mathrm{mod}\ \varpi^a)}} \frac{\rho_{kb}(\zeta)^2}{\rho_a(\xi)^2} V(\zeta,\eta), \tag{5.5}$$

in which we write

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

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

$$V(\zeta,\eta) \leqslant V_1^{1-k/s}V_2^{k/s}, \tag{5.6}$$

where

$$V_1 = \oint |\mathfrak{F}_b(\boldsymbol{\alpha};\eta)|^{2u+2}\,\mathrm{d}\boldsymbol{\alpha}$$

and

$$V_2 = \oint |\mathfrak{F}_b(\boldsymbol{\alpha};\eta)^2\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta)^{2s}|\,\mathrm{d}\boldsymbol{\alpha}.$$

By Lemma 3.2, one has $V_1 \ll (X/M^b)^{\lambda+\delta}$. On the other hand, we find from (2.14) that $V_2 = I_{b,kb}(X;\eta,\zeta)$. We therefore deduce from (5.5) and (5.6) via another application of Hölder's inequality that

$$K_{a,b}(X;\xi,\eta) \ll \rho_a(\xi)^{-2}M^{\frac{1}{2}k(k-1)(a+b)}W_1^{1-k/s}W_2^{k/s},$$

where

$$W_1 = \sum_{\substack{1\leqslant \zeta\leqslant \varpi^{kb}\\ \zeta\equiv \xi \ (\mathrm{mod}\ \varpi^a)}} \rho_{kb}(\zeta)^2(X/M^b)^{\lambda+\delta} = \rho_a(\xi)^2(X/M^b)^{\lambda+\delta}$$

and

$$W_2 = \sum_{\substack{1\leqslant \zeta\leqslant \varpi^{kb}\\ \zeta\equiv \xi \ (\mathrm{mod}\ \varpi^a)}} \rho_{kb}(\zeta)^2 I_{b,kb}(X;\eta,\zeta).$$

Yet another application of Hölder's inequality yields

$$\sum_{\xi=1}^{\varpi^a}\sum_{\eta=1}^{\varpi^b}\rho_a(\xi)^2\rho_b(\eta)^2K_{a,b}(X;\xi,\eta)\ll M^{\frac{1}{2}k(k-1)(a+b)}Z_1^{1-k/s}Z_2^{k/s},\tag{5.7}$$

where

$$Z_1=\sum_{\xi=1}^{\varpi^a}\sum_{\eta=1}^{\varpi^b}\rho_a(\xi)^2\rho_b(\eta)^2(X/M^b)^{\lambda+\delta}=\rho_0(1)^4(X/M^b)^{\lambda+\delta},$$

and, in view of (2.17),

$$\begin{aligned}Z_2&=\sum_{\eta=1}^{\varpi^b}\rho_b(\eta)^2\sum_{\xi=1}^{\varpi^a}\sum_{\substack{1\leqslant\zeta\leqslant\varpi^{kb}\\ \zeta\equiv\xi\ (\mathrm{mod}\ \varpi^a)}}\rho_{kb}(\zeta)^2I_{b,kb}(X;\eta,\zeta)\\&=\sum_{\eta=1}^{\varpi^b}\sum_{\zeta=1}^{\varpi^{kb}}\rho_b(\eta)^2\rho_{kb}(\zeta)^2I_{b,kb}(X;\eta,\zeta)=\rho_0(1)^4I_{b,kb}(X).\end{aligned}$$

An additional reference to (2.17) therefore conveys us from (5.7) to the bound

$$K_{a,b}(X)\ll\rho_0(1)^{-4}M^{\frac{1}{2}k(k-1)(b+a)}\left(\rho_0(1)^4I_{b,kb}(X)\right)^{k/s}\left(\rho_0(1)^4(X/M^b)^{\lambda+\delta}\right)^{1-k/s},$$

and the proof of the lemma is complete. □

Finally, by combining the conclusion of Lemma 4.2 with Lemma 5.2, we obtain the key bound for our iterative process.

Lemma 5.3. *Suppose that a and b are integers with $0\leqslant a<b\leqslant(12k\theta)^{-1}$. Put $H=4(k-1)b$. Then there exists an integer h, with $0\leqslant h<H$, having the property that*

$$[[K_{a,b}(X)]]\ll X^\delta M^{-kh}[[K_{b,kb+h}(X)]]^{k/s}(X/M^b)^{\Lambda(1-k/s)}+M^{-kH/6}(X/M^b)^\Lambda.$$

Proof. Recall the definitions (2.25) and (2.26). Then it follows from Lemma 5.2 that

$$[[K_{a,b}(X)]]\ll X^\delta M^\omega[[I_{b,kb}(X)]]^{k/s}(X/M^b)^{\Lambda(1-k/s)},\tag{5.8}$$

in which we have written

$$\omega=\tfrac{1}{2}k(k-1)(b+a)+(k-\tfrac{1}{2}k(k+1))(a-b)-k(k-1)b.$$

Since $\omega=0$, we may proceed to apply Lemma 4.2. Thus, we deduce that there exists an integer h with $0\leqslant h<H$ having the property that

$$I_{b,kb}(X)\ll K_{b,kb+h}(X)+M^{-sH/4}(X/M^{kb})^s(X/M^b)^{k-k(k+1)/2+\Lambda}.$$

Again referring to (2.25) and (2.26), we see that

$$[[I_{b,kb}(X)]]\ll M^{-sh}[[K_{b,kb+h}(X)]]+(X/M^b)^\Lambda M^{-sH/4}.$$

Substituting this estimate into (5.8), we obtain the bound

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

Since $M\geqslant X^{2\delta}$, the conclusion of the lemma now follows. □

6. The iterative process

The skeleton of our iterative process is now visible. Lemma 4.4 bounds $U(X)$ in terms of $K_{0,1+h}(X)$, whilst Lemma 5.3 bounds $K_{a,b}(X)$ in terms of $K_{b,kb+h}(X)$. Thus we may obtain a sequence of bounds for $U(X)$ in terms of auxiliary mean values $K_{a_n,b_n}(X)$, with a_n and b_n increasing with n. At any point in this iteration, we may apply the trivial bound for $K_{a,b}(X)$ supplied by Lemma 3.4. It transpires that, with an appropriate choice of parameters (already selected within our argument), we arrive at a contradiction whenever $\Lambda > 0$. We begin by distilling Lemma 5.3 into a more portable form.

Lemma 6.1. *Suppose that $\Lambda \geqslant 0$. Let $a, b \in \mathbb{Z}$ satisfy $0 \leqslant a < b \leqslant (12k\theta)^{-1}$. In addition, suppose that there are real numbers ψ, c and γ, with*

$$0 \leqslant c \leqslant (2\delta)^{-1}\theta, \quad \gamma \geqslant 0 \quad \text{and} \quad \psi \geqslant 0,$$

such that

$$X^\Lambda M^{\Lambda\psi} \ll X^{c\delta} M^{-\gamma} [[K_{a,b}(X)]]. \tag{6.1}$$

Then, for some integer h with $0 \leqslant h \leqslant 4kb$, one has

$$X^\Lambda M^{\Lambda\psi'} \ll X^{c'\delta} M^{-\gamma'} [[K_{b,kb+h}(X)]],$$

where

$$\psi' = (s/k)\psi + (s/k-1)b, \quad c' = (s/k)(c+1), \quad \gamma' = (s/k)\gamma + sh.$$

Proof. By hypothesis, we have $X^{c\delta} \leqslant M$. We therefore deduce from Lemma 5.3 that there exists an integer h with $0 \leqslant h < 4kb$ having the property that

$$[[K_{a,b}(X)]] \ll X^\delta M^{-kh} [[K_{b,kb+h}(X)]]^{k/s} (X/M^b)^{\Lambda(1-k/s)} + M^{-kH/6}(X/M^b)^\Lambda.$$

The hypothesised bound (6.1) consequently leads to the estimate

$$X^\Lambda M^{\Lambda\psi} \ll X^{(c+1)\delta} M^{-\gamma-kh} [[K_{b,kb+h}(X)]]^{k/s} (X/M^b)^{\Lambda(1-k/s)} + M^{-kH/6}(X/M^b)^\Lambda,$$

whence

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

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

We now come to the main act.

Lemma 6.2. *One has $\Lambda \leqslant 0$.*

Proof. Suppose first that we are able to establish the conclusion of the lemma for $s = k^2$. Then, applying the trivial estimate (2.5), we find that whenever $s > k^2$, one has

$$U_{s+k}(X;\mathfrak{a}) \leqslant \Big(\sup_{\boldsymbol{\alpha}\in[0,1)^k} |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)| \Big)^{2(s-k^2)} \oint |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)|^{2k(k+1)} \,\mathrm{d}\boldsymbol{\alpha}$$
$$\ll X^{s-k^2} U_{k(k+1)}(X;\mathfrak{a}) \ll X^{s-k^2} \left(X^{k(k+1)/2+\varepsilon} \right).$$

Hence

$$U_{s+k}(X;\mathfrak{a}) \ll X^{s+k-k(k+1)/2+\varepsilon},$$

and we find that $\Lambda \leqslant 0$. We are therefore at liberty to assume in what follows that $s = k^2$.

We may now suppose that $s = k^2$ and $\Lambda \geqslant 0$, for otherwise there is nothing to prove. We begin by applying Lemma 4.4. Thus, in view of (2.24) and (2.26), there exists an integer $h_{-1} \in \{0, 1, 2, 3\}$ such that

$$[[U(X)]] \ll (M^{h_{-1}})^{-s}[[K_{0,1+h_{-1}}(X)]].$$

We therefore deduce from (2.27) that

$$X^{\Lambda} \ll X^{\delta}[[U(X)]] \ll X^{\delta} M^{-sh_{-1}}[[K_{0,1+h_{-1}}(X)]]. \tag{6.2}$$

Next we define the sequences (h_n), (a_n), (b_n), (c_n), (ψ_n) and (γ_n) for $0 \leqslant n \leqslant N$, in such a way that

$$0 \leqslant h_{n-1} \leqslant 12kb_{n-1}, \tag{6.3}$$

and

$$X^{\Lambda} M^{\Lambda\psi_n} \ll X^{c_n\delta} M^{-\gamma_n}[[K_{a_n,b_n}(X)]]. \tag{6.4}$$

Given a fixed choice for the sequence (h_n), the remaining sequences are defined by means of the relations

$$a_{n+1} = b_n, \quad b_{n+1} = kb_n + h_n, \tag{6.5}$$

$$c_{n+1} = k(c_n + 1), \tag{6.6}$$

$$\psi_{n+1} = k\psi_n + (k-1)b_n, \tag{6.7}$$

$$\gamma_{n+1} = k\gamma_n + sh_n. \tag{6.8}$$

We put

$$a_0 = 0, \quad b_{-1} = 1, \quad b_0 = 1 + h_{-1},$$
$$\psi_0 = 0, \quad c_0 = 1, \quad \gamma_0 = sh_{-1},$$

so that both (6.3) and (6.4) hold with $n = 0$ as a consequence of our initial choice of h_{-1} together with (6.2). We prove by induction that for each non-negative integer n with $n < N$, the sequence $(h_m)_{m=-1}^{n}$ may be chosen in such a way that

$$0 \leqslant a_n \leqslant b_n \leqslant (12k\theta)^{-1}, \quad b_n \geqslant ka_n, \tag{6.9}$$

$$\psi_n \geqslant 0, \quad \gamma_n \geqslant 0, \quad 0 \leqslant c_n \leqslant (2\delta)^{-1}\theta, \tag{6.10}$$

and so that (6.3) and (6.4) both hold with n replaced by $n + 1$.

Let $0 \leqslant n \leqslant N$, and suppose that (6.3) and (6.4) both hold for the index n. We have already shown such to be the case for $n = 0$. We observe first that the relation (6.5) demonstrates that $b_m \geqslant ka_m$ for all m. Also, from (6.3) and (6.5) one finds that $b_{m+1} \leqslant 13kb_m$, whence

$$b_n \leqslant (13k)^n b_0 \leqslant 4(13k)^n.$$

Thus we see from (2.8) that $b_n \leqslant (12k\theta)^{-1}$. Also, from (6.6), (6.7) and (6.8) we see that c_n, ψ_n and γ_n are non-negative for each n. Further, we have

$$c_m \leqslant k^m + k\Big(\frac{k^m - 1}{k - 1}\Big) \leqslant 3k^m \quad (m \geqslant 0). \tag{6.11}$$

In order to bound γ_n, we begin by noting from (6.5) that

$$h_m = b_{m+1} - kb_m \quad \text{and} \quad a_m = b_{m-1}.$$

Then it follows from (6.8) that

$$\gamma_{m+1} - sb_{m+1} = k(\gamma_m - sb_m).$$

By iterating this relation, we deduce that for $m \geqslant 1$, one has

$$\gamma_m = sb_m + k^m(\gamma_0 - sb_0) = s(b_m - k^m). \tag{6.12}$$

We may now suppose that (6.4), (6.9) and (6.10) hold for the index n. An application of Lemma 6.1 therefore reveals that there exists an integer h_n satisfying (6.3) with n replaced by $n+1$, for which the upper bound (6.4) holds, also with n replaced by $n+1$. This completes the inductive step, so that (6.4) is now known to hold for $0 \leqslant n \leqslant N$.

We now exploit the bound (6.4) with $n = N$. Since $b_N \leqslant 4(13k)^N \leqslant (3\theta)^{-1}$, one finds from Lemma 3.4 that

$$[[K_{a_N,b_N}(X)]] \ll X^{\Lambda+\delta}(M^{b_N-b_{N-1}})^{k(k+1)/2}.$$

By combining this bound with (6.4) and (6.12), we obtain the estimate

$$\begin{aligned} X^\Lambda M^{\Lambda\psi_N} &\ll X^{\Lambda+(c_N+1)\delta} M^{\frac{1}{2}k(k+1)(b_N-b_{N-1})-\gamma_N} \\ &\ll X^{\Lambda+(c_N+1)\delta} M^{sk^N-(s-k(k+1)/2)b_N-\frac{1}{2}k(k+1)b_{N-1}} \\ &\ll X^{\Lambda+(c_N+1)\delta} M^{sk^N}. \end{aligned}$$

Meanwhile, from (6.11) and (2.8) we have $X^{(c_N+1)\delta} < M$. We therefore deduce that

$$\Lambda\psi_N \leqslant sk^N + 1. \tag{6.13}$$

Next, recalling that $b_m \geqslant k^m$ for each m, we deduce from (6.7) that

$$\psi_{n+1} \geqslant k\psi_n + (k-1)k^n \quad (0 \leqslant n < N),$$

whence $\psi_N \geqslant N(k-1)k^{N-1}$. We thus conclude from (6.13) that

$$\Lambda \leqslant \frac{sk^N + 1}{N(k-1)k^{N-1}} \leqslant \frac{4s}{N}.$$

Since we are at liberty to take N as large as we please in terms of s and k, we are forced to conclude that $\Lambda \leqslant 0$. In view of the discussion in the initial paragraph of the proof, this completes the proof of the lemma. □

The proof of the first estimate of Theorem 1.1 now follows. For the conclusion of Lemma 6.2 implies that when $s \geqslant k^2$, the bound

$$U_{s+k}(X; \mathfrak{a}) \ll X^{s+k-k(k+1)/2+\Lambda+\varepsilon}$$

holds with $\Lambda = 0$, and any $\varepsilon > 0$. Thus we deduce from (2.2) and (2.4) that whenever $s \geqslant k(k+1)$, then

$$\oint |f_{\mathfrak{a}}(\boldsymbol{\alpha}; X)|^{2s}\, \mathrm{d}\boldsymbol{\alpha} \ll X^{s-k(k+1)/2+\varepsilon}\Bigg(\sum_{|n|\leqslant X} |\mathfrak{a}_n|^2\Bigg)^s.$$

It remains for us to establish the second (ε-free) estimate claimed by Theorem 1.1 when $s > k(k+1)$. This task we defer to the next section.

7. The Keil-Zhao device

We now describe a relatively cheap method of slightly sharpening the first estimate of Theorem 1.1 when excess variables are present. This is motivated by recent work of Lilu Zhao [26, equation (3.10)] and Keil [14, page 608]. Kevin Hughes has also established such an ε-removal lemma by an alternate and earlier route that has priority in this topic (see [13], and also [9] for independent work on this topic). We include this section partly to highlight the utility of the Keil-Zhao device, and also to present a relatively self-contained account of the proof of Theorem 1.1. The first results of this type are due to Bourgain [3, 4], and concern quadratic problems. In view of the latter work, there is no loss in supposing throughout that $k \geqslant 3$.

The basic approach utilises the Hardy-Littlewood method. We therefore begin with some infrastructure. Write $L = X^{1/(2k)}$. Then, when $1 \leqslant q \leqslant L$, $1 \leqslant a_j \leqslant q$ $(1 \leqslant j \leqslant k)$ and $(q, a_1, \dots, a_k) = 1$, define the major arc $\mathfrak{M}(q, \mathbf{a})$ by

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

The arcs $\mathfrak{M}(q, \mathbf{a})$ are disjoint, as is easily verified. Let $\mathfrak{M}$ denote their union, and put $\mathfrak{m} = [0,1)^k \setminus \mathfrak{M}$.

Write

$$F(\boldsymbol{\alpha}; X) = \sum_{|n| \leqslant X} e(\alpha_1 n + \ldots + \alpha_k n^k).$$

Also, 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

$$S(q, \mathbf{a}) = \sum_{r=1}^{q} e((a_1 r + \ldots + a_k r^k)/q)$$

and

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

We then 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.

We make use of two basic estimates. The first follows from the argument of [21, §9] with only trivial modifications, and shows that

$$\sup_{\boldsymbol{\alpha} \in \mathfrak{m}} |F(\boldsymbol{\alpha}; X)| \ll X^{1-\tau+\varepsilon}, \tag{7.1}$$

where $\tau^{-1} = 4k^2$. The second estimate we record in the shape of a lemma.

Lemma 7.1. *Suppose that $u > \frac{1}{2}k(k+1) + 2$. Then one has*

$$\int_{\mathfrak{M}} |F(\boldsymbol{\alpha}; X)|^u \, \mathrm{d}\boldsymbol{\alpha} \ll_u X^{u - k(k+1)/2}.$$

Proof. It follows from [19, Theorem 7.2] that when $\boldsymbol{\alpha} \in \mathfrak{M}(q, \mathbf{a}) \subseteq \mathfrak{M}$, one has

$$F(\boldsymbol{\alpha}; X) - V(\boldsymbol{\alpha}; q, \mathbf{a}) \ll q + X|q\alpha_1 - a_1| + \ldots + X^k|q\alpha_k - a_k| \ll L^2.$$

Thus we obtain

$$\int_{\mathfrak{M}} |F(\boldsymbol{\alpha}; X)|^u \, \mathrm{d}\boldsymbol{\alpha} \ll \int_{\mathfrak{M}} (L^2)^u \, \mathrm{d}\boldsymbol{\alpha} + \int_{\mathfrak{M}} |V(\boldsymbol{\alpha})|^u \, \mathrm{d}\boldsymbol{\alpha}. \tag{7.2}$$

But $\mathrm{mes}(\mathfrak{M}) \ll L^{2k+1}X^{-k(k+1)/2}$, and so

$$\int_{\mathfrak{M}} (L^2)^u \, \mathrm{d}\boldsymbol{\alpha} \ll L^{2u+2k+1}X^{-k(k+1)/2} \ll X^{u-k(k+1)/2}. \tag{7.3}$$

Meanwhile, one finds that

$$\int_{\mathfrak{M}} |V(\boldsymbol{\alpha})|^u \, \mathrm{d}\boldsymbol{\alpha} = \mathfrak{S}\mathfrak{J}, \tag{7.4}$$

where

$$\mathfrak{J} = \int_{\mathfrak{B}} |I(\boldsymbol{\beta}; X)|^u \, \mathrm{d}\boldsymbol{\alpha} \quad \text{and} \quad \mathfrak{S} = \sum_{1 \leqslant q \leqslant L} \sum_{\substack{1 \leqslant \mathbf{a} \leqslant q \\ (q, a_1, \ldots, a_k) = 1}} |q^{-1}S(q, \mathbf{a})|^u,$$

in which we write

$$\mathfrak{B} = [-LX^{-1}, LX^{-1}] \times \ldots \times [-LX^{-k}, LX^{-k}].$$

Since [1, Theorem 1.3] shows that the singular integral

$$\int_{\mathbb{R}^k} |I(\boldsymbol{\beta}; 1)|^{2s} \, \mathrm{d}\boldsymbol{\beta}$$

converges for $2s > \frac{1}{2}k(k+1) + 1$, we find via two changes of variables that

$$\mathfrak{J} \leqslant X^{u-k(k+1)/2} \int_{\mathbb{R}^k} |I(\boldsymbol{\beta}; 1)|^u \, \mathrm{d}\boldsymbol{\beta} \ll X^{u-k(k+1)/2}.$$

Also, by reference to [1, Theorem 2.4], one sees that the singular series

$$\sum_{q=1}^{\infty} \sum_{\substack{1 \leqslant \mathbf{a} \leqslant q \\ (q, a_1, \ldots, a_k) = 1}} |q^{-1}S(q, \mathbf{a})|^{2s}$$

converges for $2s > \frac{1}{2}k(k+1) + 2$, and hence

$$\mathfrak{S} \leqslant \sum_{q=1}^{\infty} \sum_{\substack{1 \leqslant \mathbf{a} \leqslant q \\ (q, a_1, \ldots, a_k) = 1}} |q^{-1}S(q, \mathbf{a})|^u \ll 1.$$

On substituting these estimates into (7.4), we deduce that

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

and hence the conclusion of the lemma follows from (7.2) and (7.3). □

We are now equipped to apply the Keil-Zhao device.

Lemma 7.2. *Suppose that w is a real number with $w \geqslant \frac{1}{2}k(k+1)$ for which one has the estimate*

$$\oint |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)|^{2w}\,\mathrm{d}\boldsymbol{\alpha} \ll X^{w-k(k+1)/2+\varepsilon}. \tag{7.5}$$

Then whenever $s > \max\{w, \frac{1}{2}k(k+1)+2\}$, one has

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

Proof. The hypothesis of the statement of the lemma permits us to assume that $s = w + k\nu$, for some $\nu > 0$. Let δ be a positive number sufficiently small in terms of k and ν, and put

$$\mathfrak{B} = \{\boldsymbol{\alpha} \in [0,1)^k : |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)| > X^{1/2-\delta}\}.$$

Then the upper bound (7.5) ensures that

$$\operatorname{mes}(\mathfrak{B}) \leqslant \big(X^{1/2-\delta}\big)^{-2w} \oint |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)|^{2w}\,\mathrm{d}\boldsymbol{\alpha} \ll X^{2w\delta-k(k+1)/2+\varepsilon}. \tag{7.6}$$

Putting $\mathfrak{b} = [0,1)^k \setminus \mathfrak{B}$, it follows that

$$\begin{aligned}\int_{\mathfrak{b}} |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha} &\leqslant \Big(\sup_{\boldsymbol{\alpha}\in\mathfrak{b}} |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)|\Big)^{2k\nu} \oint |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)|^{2w}\,\mathrm{d}\boldsymbol{\alpha}\\ &\ll (X^{1/2-\delta})^{2k\nu} X^{w-k(k+1)/2+\varepsilon}\\ &\ll X^{s-k(k+1)/2-\delta k\nu}.\end{aligned} \tag{7.7}$$

We next consider the mean value

$$\Upsilon = \int_{\mathfrak{B}} |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha}.$$

We first rewrite Υ in the form

$$\Upsilon = \sum_{|\mathbf{n}|\leqslant X} \mathfrak{c}(\mathbf{n}) \int_{\mathfrak{B}} |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)|^{2s-2} e(\Psi(\mathbf{n};\boldsymbol{\alpha}))\,\mathrm{d}\boldsymbol{\alpha},$$

in which

$$\mathfrak{c}(\mathbf{n}) = \rho(\mathfrak{a};X)^{-2}\mathfrak{a}_{n_1}\overline{\mathfrak{a}}_{n_2}$$

and

$$\Psi(\mathbf{n};\boldsymbol{\alpha}) = \psi(n_1;\boldsymbol{\alpha}) - \psi(n_2;\boldsymbol{\alpha}).$$

Applying Cauchy's inequality, we deduce that

$$\Upsilon^2 \leqslant \rho \sum_{|\mathbf{n}|\leqslant X} \int_{\mathfrak{B}}\int_{\mathfrak{B}} |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\beta};X)|^{2s-2} e(\Psi(\mathbf{n};\boldsymbol{\alpha}-\boldsymbol{\beta}))\,\mathrm{d}\boldsymbol{\alpha}\,\mathrm{d}\boldsymbol{\beta},$$

where

$$\rho = \rho(\mathfrak{a};X)^{-4}\Big(\sum_{|n|\leqslant X} |\mathfrak{a}_n|^2\Big)^2 = 1.$$

Thus we obtain the relation

$$\Upsilon^2 \leqslant \int_{\mathfrak{B}}\int_{\mathfrak{B}} |F(\boldsymbol{\alpha}-\boldsymbol{\beta};X)^2 \widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)^{2s-2}\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\beta};X)^{2s-2}|\,\mathrm{d}\boldsymbol{\alpha}\,\mathrm{d}\boldsymbol{\beta}. \tag{7.8}$$

Our first observation concerning the integral on the right hand side of (7.8) concerns the set of points $(\boldsymbol{\alpha},\boldsymbol{\beta})\in\mathfrak{B}\times\mathfrak{B}$ for which $\boldsymbol{\alpha}-\boldsymbol{\beta}\in\mathfrak{m}$. On applying the bound (7.1), we deduce via a trivial inequality for $\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\theta};X)$ that

$$\int_{\mathfrak{B}}\int_{\mathfrak{B}}\limits_{\boldsymbol{\alpha}-\boldsymbol{\beta}\in\mathfrak{m}}|F(\boldsymbol{\alpha}-\boldsymbol{\beta};X)^2\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)^{2s-2}\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\beta};X)^{2s-2}|\,\mathrm{d}\boldsymbol{\alpha}\,\mathrm{d}\boldsymbol{\beta}$$

$$\leqslant\Big(\sup_{\boldsymbol{\alpha}-\boldsymbol{\beta}\in\mathfrak{m}}|F(\boldsymbol{\alpha}-\boldsymbol{\beta};X)|\Big)^2(X^{1/2})^{4s-4}\int_{\mathfrak{B}}\int_{\mathfrak{B}}\mathrm{d}\boldsymbol{\alpha}\,\mathrm{d}\boldsymbol{\beta}$$
$$\ll(X^{1-\tau+\varepsilon})^2X^{2s-2}\,(\mathrm{mes}(\mathfrak{B}))^2\,.$$

On applying (7.6), therefore, we infer that

$$\int_{\mathfrak{B}}\int_{\mathfrak{B}}\limits_{\boldsymbol{\alpha}-\boldsymbol{\beta}\in\mathfrak{m}}|F(\boldsymbol{\alpha}-\boldsymbol{\beta};X)^2\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)^{2s-2}\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\beta};X)^{2s-2}|\,\mathrm{d}\boldsymbol{\alpha}\,\mathrm{d}\boldsymbol{\beta}$$
$$\ll X^{2s-2\tau+\varepsilon}(X^{2w\delta-k(k+1)/2+\varepsilon})^2$$
$$\ll X^{2s-k(k+1)-\tau}. \tag{7.9}$$

On the other hand, by applying the trivial inequality

$$|z_1\dots z_r|\leqslant|z_1|^r+\ldots+|z_r|^r,$$

it follows that

$$|\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)^{2s-2}\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\beta};X)^{2s-2}|\ll|\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)^{2s}\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\beta};X)^{2s-4}|$$
$$+|\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)^{2s-4}\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\beta};X)^{2s}|.$$

Then by symmetry, we obtain the bound

$$\int_{\mathfrak{B}}\int_{\mathfrak{B}}\limits_{\boldsymbol{\alpha}-\boldsymbol{\beta}\in\mathfrak{M}}|F(\boldsymbol{\alpha}-\boldsymbol{\beta};X)^2\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)^{2s-2}\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\beta};X)^{2s-2}|\,\mathrm{d}\boldsymbol{\alpha}\,\mathrm{d}\boldsymbol{\beta}$$
$$\ll\int_{\mathfrak{B}}\int_{\mathfrak{B}}\limits_{\boldsymbol{\alpha}-\boldsymbol{\beta}\in\mathfrak{M}}|F(\boldsymbol{\alpha}-\boldsymbol{\beta};X)^2\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)^{2s-4}\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\beta};X)^{2s}|\,\mathrm{d}\boldsymbol{\alpha}\,\mathrm{d}\boldsymbol{\beta}.$$

An application of Hölder's inequality therefore reveals that

$$\int_{\mathfrak{B}}\int_{\mathfrak{B}}\limits_{\boldsymbol{\alpha}-\boldsymbol{\beta}\in\mathfrak{M}}|F(\boldsymbol{\alpha}-\boldsymbol{\beta};X)^2\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)^{2s-2}\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\beta};X)^{2s-2}|\,\mathrm{d}\boldsymbol{\alpha}\,\mathrm{d}\boldsymbol{\beta}\ll I_1^{2/s}I_2^{1-2/s}, \tag{7.10}$$

where

$$I_1=\int_{\mathfrak{B}}\int_{\mathfrak{B}}\limits_{\boldsymbol{\alpha}-\boldsymbol{\beta}\in\mathfrak{M}}|F(\boldsymbol{\alpha}-\boldsymbol{\beta};X)^s\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\beta};X)^{2s}|\,\mathrm{d}\boldsymbol{\alpha}\,\mathrm{d}\boldsymbol{\beta}$$

and

$$I_2=\int_{\mathfrak{B}}\int_{\mathfrak{B}}\limits_{\boldsymbol{\alpha}-\boldsymbol{\beta}\in\mathfrak{M}}|\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)^{2s}\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\beta};X)^{2s}|\,\mathrm{d}\boldsymbol{\alpha}\,\mathrm{d}\boldsymbol{\beta}.$$

By Lemma 7.1, since $s > \frac{1}{2}k(k+1)+2$, we see that

$$I_1 \ll \left(\int_{\mathfrak{M}} |F(\boldsymbol{\theta};X)|^s \, \mathrm{d}\boldsymbol{\theta} \right) \left(\int_{\mathfrak{B}} |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\beta};X)|^{2s} \, \mathrm{d}\boldsymbol{\beta} \right) \ll X^{s-k(k+1)/2}\Upsilon.$$

On the other hand,

$$I_2 \leqslant \left(\int_{\mathfrak{B}} |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \right) \left(\int_{\mathfrak{B}} |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\beta};X)|^{2s} \, \mathrm{d}\boldsymbol{\beta} \right) = \Upsilon^2.$$

Thus, we infer from (7.10) that

$$\begin{aligned} \int_{\mathfrak{B}} \int_{\mathfrak{B}}_{\boldsymbol{\alpha}-\boldsymbol{\beta}\in\mathfrak{M}} & |F(\boldsymbol{\alpha}-\boldsymbol{\beta};X)^2 \widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)^{2s-2} \widetilde{f}_{\mathfrak{a}}(\boldsymbol{\beta};X)^{2s-2}| \, \mathrm{d}\boldsymbol{\alpha} \, \mathrm{d}\boldsymbol{\beta} \\ & \ll (X^{s-k(k+1)/2}\Upsilon)^{2/s} (\Upsilon^2)^{1-2/s} \\ & \ll X^{2-k(k+1)/s}\Upsilon^{2-2/s}. \end{aligned} \tag{7.11}$$

Combining (7.8) and (7.9) with (7.11), we deduce that

$$\Upsilon^2 \ll X^{2s-k(k+1)-\tau} + X^{2-k(k+1)/s}\Upsilon^{2-2/s},$$

whence

$$\int_{\mathfrak{B}} |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} = \Upsilon \ll X^{s-k(k+1)/2}.$$

Finally, we combine the last bound with (7.7), obtaining the estimate

$$\begin{aligned} \oint |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} &= \int_{\mathfrak{B}} |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} + \int_{\mathfrak{b}} |\widetilde{f}_{\mathfrak{a}}(\boldsymbol{\alpha};X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \\ &\ll X^{s-k(k+1)/2-\delta k\nu} + X^{s-k(k+1)/2} \ll X^{s-k(k+1)/2}. \end{aligned}$$

This completes the proof of the lemma. □

Since the first conclusion of Theorem 1.1, already established, confirms the validity of the hypothesis (7.5) with $w = k(k+1)$, the second conclusion of Theorem 1.1 is immediate from Lemma 7.2. Hence Theorem 1.1 has now been proved in full.

8. Some consequences of Theorem 1.1

We take some space in this section to provide brief accounts of the proofs of the two corollaries to Theorem 1.1. This is all quite standard, and so we feel justified in economising on detail.

The proof of Corollary 1.2. It follows from Theorem 1.1 that whenever $k \geqslant 2$, $p \geqslant 2k(k+1)$ and $\varepsilon > 0$, then for any complex sequence $(\mathfrak{a}_n)_{n\in\mathbb{Z}}$, one has

$$\oint |f_{\mathfrak{a}}(\boldsymbol{\alpha};N)|^p \, \mathrm{d}\boldsymbol{\alpha} \ll N^{(p-k(k+1))/2+\varepsilon} \Bigl(\sum_{|n|\leqslant N} |\mathfrak{a}_n|^2 \Bigr)^{p/2}.$$

The estimate (1.3) therefore follows on raising left and right hand sides of this relation to the power $1/p$, wherein

$$K_{p,N} \ll (N^{(p-k(k+1))/2+\varepsilon})^{1/p} = N^{(1-\theta)/2+\varepsilon}.$$

Here, one may take $\varepsilon=0$ whenever $p>2k(k+1)$. The first claim of Corollary 1.2 has therefore been established.

We now address the bound for $A_{p,N}$. Let $g:(\mathbb{R}/\mathbb{Z})^k\to\mathbb{C}$ have Fourier series defined as in (1.1), and put

$$g_0(\boldsymbol{\alpha})=\sum_{|n|\leqslant N}\hat{g}(n,n^2,\ldots,n^k)e(\alpha_1n+\ldots+\alpha_kn^k).$$

Then, by orthogonality in combination with Hölder's inequality, one has

$$\begin{aligned}\sum_{|n|\leqslant N}|\hat{g}(n,n^2,\ldots,n^k)|^2&=\oint g_0(\boldsymbol{\alpha})\tilde{g}(-\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha}\\&\leqslant\left(\oint|g_0(\boldsymbol{\alpha})|^p\,\mathrm{d}\boldsymbol{\alpha}\right)^{1/p}\left(\oint|\tilde{g}(\boldsymbol{\alpha})|^{p'}\,\mathrm{d}\boldsymbol{\alpha}\right)^{1/p'},\end{aligned}\tag{8.1}$$

in which we write $p'=p/(p-1)$. The first integral on the right hand side may be bounded via Theorem 1.1, giving

$$\oint|g_0(\boldsymbol{\alpha})|^p\,\mathrm{d}\boldsymbol{\alpha}\ll N^{(p-k(k+1))/2+\varepsilon}\Bigl(\sum_{|n|\leqslant N}|\hat{g}(n,n^2,\ldots,n^k)|^2\Bigr)^{p/2}.$$

On substituting this bound into (8.1), we obtain the relation

$$\sum_{|n|\leqslant N}|\hat{g}(n,n^2,\ldots,n^k)|^2\ll N^{(1-\theta)/2+\varepsilon}\Bigl(\sum_{|n|\leqslant N}|\hat{g}(n,n^2,\ldots,n^k)|^2\Bigr)^{1/2}\|\tilde{g}\|_{p'}.$$

On disentangling this relation, and noting that $\|\tilde{g}\|_{p'}=\|g\|_{p'}$, we obtain the bound

$$\sum_{|n|\leqslant N}|\hat{g}(n,n^2,\ldots,n^k)|^2\ll A_{p,N}\|g\|_{p'}^2,$$

with $A_{p,N}\ll N^{1-\theta+\varepsilon}$. The second conclusion of Corollary 1.2 follows, noting that one may take $\varepsilon=0$ when $p>2k(k+1)$. ☐

The proof of Corollary 1.3. Consider positive integers $k_1,\ldots,k_t$ with $1\leqslant k_1<k_2<\ldots<k_t=k$, and denote by $l_1,\ldots,l_u$ those positive integers with $1\leqslant l_1<l_2<\ldots<l_u<k$ for which

$$\{k_1,\ldots,k_t\}\cup\{l_1,\ldots,l_u\}=\{1,2,\ldots,k\}.$$

Plainly, therefore, one has $u=k-t$, and moreover,

$$\sum_{i=1}^u l_i=\sum_{j=1}^k j-\sum_{m=1}^t k_m=\tfrac{1}{2}k(k+1)-K.\tag{8.2}$$

By orthogonality, the mean value

$$I=\oint\Bigl|\sum_{|n|\leqslant N}\mathfrak{a}_ne(\alpha_1n^{k_1}+\ldots+\alpha_tn^{k_t})\Bigr|^{2s}\,\mathrm{d}\boldsymbol{\alpha}$$

counts the number of integral solutions of the system of equations

$$\sum_{i=1}^s (x_i^{k_m} - y_i^{k_m}) = 0 \quad (1 \leqslant m \leqslant t),$$

with $|\mathbf{x}|, |\mathbf{y}| \leqslant N$, with each solution counted with weight

$$\prod_{i=1}^s \mathfrak{a}_{x_i} \overline{\mathfrak{a}}_{y_i}. \tag{8.3}$$

Given any one solution $\mathbf{x}, \mathbf{y}$ counted by I, there exist integers h_j $(1 \leqslant j \leqslant k)$ for which

$$\sum_{i=1}^s (x_i^j - y_i^j) = h_j \quad (1 \leqslant j \leqslant k). \tag{8.4}$$

Indeed, one has $h_j = 0$ whenever $j = k_m$ for some suffix m. It is evident, moreover, that $|h_j| \leqslant 2sN^j$ $(1 \leqslant j \leqslant k)$.

When $\mathbf{h} \in \mathbb{Z}^k$, consider the integral

$$I(\mathbf{h}) = \oint |f_{\mathfrak{a}}(\boldsymbol{\beta}; N)|^{2s} e(-h_1\beta_1 - \ldots - h_k\beta_k)\, \mathrm{d}\boldsymbol{\beta}.$$

By orthogonality, this integral counts the integral solutions of the system of equations (8.4) with $|\mathbf{x}|, |\mathbf{y}| \leqslant N$, and with each solution counted with weight (8.3). Thus we see that

$$I = \sum_{|h_{l_1}| \leqslant 2sN^{l_1}} \cdots \sum_{|h_{l_u}| \leqslant 2sN^{l_u}} I(\mathbf{h}),$$

in which we put $h_j = 0$ whenever $j = k_m$. By the triangle inequality, it therefore follows that

$$\begin{aligned} I &\leqslant \sum_{|h_{l_1}| \leqslant 2sN^{l_1}} \cdots \sum_{|h_{l_u}| \leqslant 2sN^{l_u}} I(\mathbf{0}) \\ &\ll N^{l_1 + \ldots + l_u} \oint |f_{\mathfrak{a}}(\boldsymbol{\beta}; N)|^{2s}\, \mathrm{d}\boldsymbol{\beta}. \end{aligned}$$

Thus we deduce from (8.2) and Theorem 1.1 that when $s \geqslant k(k+1)$, one has

$$\begin{aligned} I &\ll N^{\frac{1}{2}k(k+1) - K} \left(N^{s - \frac{1}{2}k(k+1) + \varepsilon} \Bigl(\sum_{|n| \leqslant N} |\mathfrak{a}_n|^2 \Bigr)^s \right) \\ &\ll N^{s - K + \varepsilon} \Bigl(\sum_{|n| \leqslant N} |\mathfrak{a}_n|^2 \Bigr)^s. \end{aligned}$$

Here, one may take $\varepsilon = 0$ when $s > k(k+1)$. This completes the proof of Corollary 1.3. □

9. Wider applications of weighted efficient congruencing

As we have already noted in the introduction, the ideas of this paper can be transported to deliver weighted variants of any mean value estimate established via efficient congruencing methods. Indeed, the basic [7, 22] and multigrade variants of efficient congruencing introduced in [23, 24, 25] may all be modified to accommodate the weighted setting appropriate for restriction theory. Although this task is not especially easy, by adapting these methods one may establish the Main Conjecture (recorded above as Conjecture 1.4) for

$$1 \leqslant s \leqslant \tfrac{1}{2}k(k+1) - (\tfrac{1}{3} + o(1))k \quad (k \text{ large})$$

and

$$s \geqslant k(k-1) \quad (k \geqslant 3).$$

Moreover, one may take $\varepsilon = 0$ when $s > k(k-1)$. The last result, in particular, confirms the Main Conjecture in full for $k = 3$. As we have noted in the introduction, the associated arguments are of sufficient complexity that, were we to establish them in full as the main thrust of this paper, we would obscure the basic principles of the weighted efficient congruencing method. Instead, we intend to provide complete accounts of the proofs of these conclusions as special cases of more general results in subsequent papers.

Such ideas also extend to multidimensional settings. Consider, for example, a system of polynomials $\mathbf{F} = (F_1, \dots, F_r)$, with $F_i \in \mathbb{Z}[x_1, \dots, x_d]$ $(1 \leqslant i \leqslant r)$. Following [16, §2], we say that $\mathbf{F}$ is translation-dilation invariant if:
(i) the polynomials $F_1, \dots, F_r$ are each homogeneous of positive degree, and
(ii) there exist polynomials

$$c_{jl} \in \mathbb{Z}[\xi_1, \dots, \xi_d] \quad (1 \leqslant j \leqslant r \text{ and } 0 \leqslant l \leqslant j),$$

with $c_{jj} = 1$ for $1 \leqslant j \leqslant r$, having the property that whenever $\boldsymbol{\xi} \in \mathbb{Z}^d$, then

$$F_j(\mathbf{x} + \boldsymbol{\xi}) = c_{j0}(\boldsymbol{\xi}) + \sum_{l=1}^{j} c_{jl}(\boldsymbol{\xi}) F_l(\mathbf{x}) \quad (1 \leqslant j \leqslant r).$$

It follows that the system

$$\sum_{i=1}^{s} (F_j(\mathbf{x}_i) - F_j(\mathbf{y}_i)) = 0 \quad (1 \leqslant j \leqslant r), \tag{9.1}$$

possesses an integral solution $\mathbf{x}, \mathbf{y}$ if and only if, for each $\boldsymbol{\xi} \in \mathbb{Z}^d$ and $\lambda \in \mathbb{Z} \setminus \{0\}$, one has

$$\sum_{i=1}^{s} (F_j(\lambda \mathbf{x}_i + \boldsymbol{\xi}) - F_j(\lambda \mathbf{y}_i + \boldsymbol{\xi})) = 0 \quad (1 \leqslant j \leqslant r),$$

whence the system (9.1) is translation-dilation invariant. Such systems are easily generated by taking one or more seed polynomials $G(\mathbf{x})$, and then appending to the system the successive partial derivatives with respect to each variable. Without loss, one may then consider only reduced systems $\mathbf{F}$ in which $F_1, \dots, F_r$ are linearly independent over $\mathbb{Q}$.

With the above system $\mathbf{F}$, we introduce some parameters in order to ease subsequent discussion. We refer to the number of variables $d = d(\mathbf{F})$ in $\mathbf{F}$ as the *dimension* of the system, the number of forms $r = r(\mathbf{F})$ comprising $\mathbf{F}$ as its *rank*, and we denote by $k_j = k_j(\mathbf{F})$ the total degree of F_j. Finally, the *degree* $k = k(\mathbf{F})$ of the system is defined by

$$k(\mathbf{F}) = \max_{1\leqslant j\leqslant r} k_j(\mathbf{F}),$$

and the *weight* $K = K(\mathbf{F})$ of the system is

$$K(\mathbf{F}) = \sum_{j=1}^{r} k_j(\mathbf{F}).$$

Then by adapting the methods of this paper in a pedestrian manner within the arguments of [16], one may establish the following.

Theorem 9.1. *Let $\mathbf{F}$ be a reduced translation-dilation invariant system of polynomials having dimension d, rank r, degree k and weight K. Suppose that s is a natural number with $s \geqslant r(k+1)$. Then for each $\varepsilon > 0$, and any complex sequence $(\mathfrak{a}_\mathbf{n})_{\mathbf{n}\in\mathbb{Z}^d}$, one has*

$$\oint \Bigl| \sum_{\mathbf{n}\in[-N,N]^d} \mathfrak{a}_\mathbf{n} e(\alpha_1 F_1(\mathbf{n}) + \ldots + \alpha_r F_r(\mathbf{n})) \Bigr|^{2s} \mathrm{d}\boldsymbol{\alpha} \ll N^{sd-K+\varepsilon} \Bigl(\sum_{\mathbf{n}\in[-N,N]^d} |\mathfrak{a}_\mathbf{n}|^2 \Bigr)^s .$$

Moreover, one may take $\varepsilon = 0$ when $s > r(k+1)$.

The arguments required for the proof of Theorem 9.1 are straightforward analogues of those required in the case $d = 1$ central to this paper, and involve none of the complications demanded by the multigrade efficient congruencing methods of [23, 24, 25]. We consequently propose to expand no further on this subject, leaving the reader to complete the routine exercises needed for its proof, and to apply [16] as the necessary framework.

References

[1] G. I. Arkhipov, V. N. Chubarikov and A. A. Karatsuba, *Trigonometric sums in number theory and analysis*, de Gruyter Expositions in Mathematics, **39**, Walter de Gruyter, Berlin, 2004.

[2] J. Bennett, A. Carbery and T. Tao, *On the multilinear restriction and Kakeya conjectures*, Acta Math. **196** (2006), no. 2, 261–302.

[3] J. Bourgain, *On $\Lambda(p)$-subsets of squares*, Israel J. Math. **67** (1989), no. 3, 291–311.

[4] J. Bourgain, *Fourier transform restriction phenomena for certain lattice subsets and applications to non-linear evolution equations, Part I: Schrödinger equations*, Geom. Funct. Anal. **3** (1993), no. 2, 107–156.

[5] J. Bourgain, *Fourier transform restriction phenomena for certain lattice subsets and applications to non-linear evolution equations, Part II: The KdV-equation*, Geom. Funct. Anal. **3** (1993), no. 3, 209–262.

[6] J. Bourgain and C. Demeter, *Decouplings for curves and hypersurfaces with non-zero Gaussian curvature*, preprint available as arXiv:1409.1634.

[7] K. Ford and T. D. Wooley, *On Vinogradov's mean value theorem: strongly diagonal behaviour via efficient congruencing*, Acta Math. **213** (2014), 199–236.

[8] K. Henriot, *Logarithmic bounds for translation-invariant equations in squares*, Internat. Math. Res. Notices, to appear.

[9] K. Henriot, *Discrete restriction, ε-removal, and additive equations in dense variables*, in preparation.

[10] Y. Hu and X. Li, *Discrete Fourier restriction associated with KdV equations*, Anal. PDE **6** (2013), no. 4, 859–892.

[11] Y. Hu and X. Li, *Discrete Fourier restriction associated with Schrödinger equations*, Rev. Mat. Iberoam. **30** (2014), no. 4, 1281–1300.

[12] L.-K. Hua, *An improvement of Vinogradov's mean value theorem and several applications*, Quart. J. Math. Oxford **20** (1949), 48–61.

[13] K. Hughes, *Extending Bourgain's discrete epsilon-removal lemma*, in preparation.

[14] E. Keil, *On a diagonal quadric in dense variables*, Glasgow Math. J. **56** (2014), no. 3, 601–628.

[15] E. Keil, *Some refinements for translation invariant quadratic forms in dense sets*, preprint available as arXiv:1408.1535.

[16] S. T. Parsell, S. M. Prendiville and T. D. Wooley, *Near-optimal mean value estimates for multidimensional Weyl sums*, Geom. Funct. Anal. **23** (2013), 1962–2024.

[17] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970.

[18] T. Tao, *Some recent progress on the restriction conjecture*, Fourier analysis and convexity, 217–243, Appl. Numer. Harmon. Anal., Birkhäuser Boston, Boston, MA, 2004.

[19] R. C. Vaughan, *The Hardy-Littlewood method*, Cambridge University Press, Cambridge, 1997.

[20] I. M. Vinogradov, *The method of trigonometrical sums in the theory of numbers*, Trav. Inst. Math. Stekloff **23** (1947), 109pp.

[21] T. D. Wooley, *Vinogradov's mean value theorem via efficient congruencing*, Annals of Math. (2) **175** (2012), no. 3, 1575–1627.

[22] T. D. Wooley, *Vinogradov's mean value theorem via efficient congruencing, II*, Duke Math. J. **162** (2013), no. 4, 673–730.

[23] T. D. Wooley, *Multigrade efficient congruencing and Vinogradov's mean value theorem*, Proc. London Math. Soc. (in press).

[24] T. D. Wooley, *Approximating the main conjecture in Vinogradov's mean value theorem*, submitted; arXiv:1401.2932.

[25] T. D. Wooley, *The cubic case of the main conjecture in Vinogradov's mean value theorem*, submitted; arXiv:1401.3150.

[26] Lilu Zhao, *On the Waring-Goldbach problem for fourth and sixth powers*, Proc. London Math. Soc. (3) **108** (2014), no. 6, 1593–1622.

School of Mathematics, University of Bristol, University Walk, Clifton, Bristol BS8 1TW, United Kingdom

E-mail address: matdw@bristol.ac.uk