VINOGRADOV'S MEAN VALUE THEOREM VIA EFFICIENT CONGRUENCING, II

TREVOR D. WOOLEY*

Abstract. We apply the efficient congruencing method to estimate Vinogradov's integral for moments of order $2s$, with $1 \leqslant s \leqslant k^2 - 1$. Thereby, we show that quasi-diagonal behaviour holds when $s = o(k^2)$, we obtain near-optimal estimates for $1 \leqslant s \leqslant \frac{1}{4}k^2 + k$, and optimal estimates for $s \geqslant k^2 - 1$. In this way we come half way to proving the main conjecture in two different directions. There are consequences for estimates of Weyl type, and in several allied applications. Thus, for example, the anticipated asymptotic formula in Waring's problem is established for sums of s kth powers of natural numbers whenever $s \geqslant 2k^2 - 2k - 8$ $(k \geqslant 6)$.

1. Introduction

Estimates stemming from Vinogradov's mean value theorem deliver bounds for exponential sums of large degree, both in mean and pointwise, beyond the competence of alternate approaches. The ubiquity of such exponential sums in analytic number theory, in the analysis for example of the Riemann zeta function, in Waring's problem, and beyond, accounts for the high profile of Vinogradov's methods in the associated literature. In recent work, we established a version of Vinogradov's mean value theorem which achieves an essentially optimal upper bound with a number of variables only twice the number conjectured to be best possible (see [20]). For systems of degree k, previous estimates missed such a bound by a factor of order $\log k$. Our earlier approach provides no upper bounds when the number of variables is smaller, precluding the possibility of applications involving the finer features of these mean values. Our goal in this paper is to remedy this deficiency, at the same time strengthening our previous conclusions. It transpires that we are able to come within a hair's breadth of proving the main conjecture concerning Vinogradov's mean value theorem in half of the basic interval of relevant moments. Such developments illustrate the flexibility of the new efficient congruencing method introduced in [20].

We now introduce some notation. When $k \in \mathbb{N}$ and $\boldsymbol{\alpha} \in \mathbb{R}^k$, define

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

2010 *Mathematics Subject Classification.* 11L15, 11L07, 11P05, 11P55.

Key words and phrases. Exponential sums, Waring's problem, Hardy-Littlewood method.

*Supported by a Royal Society Wolfson Research Merit Award.

where $e(z)$ denotes $e^{2\pi iz}$. Our goal is to estimate the mean value

$$J_{s,k}(X)=\oint |f_k(\boldsymbol{\alpha};X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha},$$

which by orthogonality counts the solutions of the Diophantine system

$$x_1^j+\ldots+x_s^j=y_1^j+\ldots+y_s^j\quad (1\leqslant j\leqslant k),$$

with $1\leqslant \mathbf{x},\mathbf{y}\leqslant X$. Here and elsewhere, we employ the convention that whenever $G:[0,1)^k\rightarrow \mathbb{C}$ is integrable, then

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

In addition, we make slightly unconventional use of vector notation. Thus, for example, we write $1\leqslant \mathbf{x}\leqslant X$ to denote that $1\leqslant x_i\leqslant X$ $(1\leqslant i\leqslant s)$.

We complete the proof of our basic estimate for $J_{s,k}(X)$ in §8. Here and elsewhere, so far as implicit constants associated with Vinogradov's notation $\ll$ and $\gg$ are concerned, we suppress mention of dependence on s, k and ε.

Theorem 1.1. *Suppose that s and k are natural numbers with $k\geqslant 3$ and $s\geqslant k^2-1$. Then, for each $\varepsilon>0$, one has $J_{s,k}(X)\ll X^{2s-\frac{1}{2}k(k+1)+\varepsilon}$.*

Prior to the introduction of the efficient congruencing method, conclusions of the type supplied by Theorem 1.1 were available only for $s\geqslant (1+o(1))k^2\log k$ (see [1], [15], [16], [18] and earlier work of Hua [8]). In [20, Theorem 1.1], meanwhile, we showed that $J_{s,k}(X)\ll X^{2s-\frac{1}{2}k(k+1)+\varepsilon}$ for $s\geqslant k(k+1)$, and this yields the conclusion of Theorem 1.1 with the condition $s\geqslant k^2-1$ replaced by $s\geqslant k^2+k$. Our new result is consequently rather sharper than that of [20], which in terms of the constraint on the number of variables already comes within a factor 2 of the widely held conjecture that $J_{s,k}(X)\ll X^{2s-\frac{1}{2}k(k+1)+\varepsilon}$ for $s\geqslant \frac{1}{2}k(k+1)$.

There are numerous consequences of Theorem 1.1, with refinements available for estimates of Weyl sums, fractional parts of polynomials, and various Diophantine problems. Since these improvements are modest in scale compared to those made available in our previous work [20], we defer discussion of the bulk of such matters to §11. For the moment, we choose instead to pursue the more subtle features of the behaviour of the mean value $J_{s,k}(X)$.

In order to motivate a discussion of the mean value $J_{s,k}(X)$ for smaller values of s, we begin by recalling the lower bound

$$J_{s,k}(X)\gg X^s+X^{2s-\frac{1}{2}k(k+1)}.\tag{1.1}$$

The closely associated conjectural upper bound

$$J_{s,k}(X)\ll X^{\varepsilon}(X^s+X^{2s-\frac{1}{2}k(k+1)}).\tag{1.2}$$

is approximated for $0<s\leqslant \frac{1}{2}k(k+1)$ by an estimate of the shape

$$J_{s,k}(X)\ll X^{s+\delta_{s,k}+\varepsilon},\tag{1.3}$$

provided that $\delta_{s,k}\geqslant 0$ is small. Suppose that (1.3) holds for an exponent sequence $\delta_{s,k}$ with $\delta_{s,k}\rightarrow 0$ as $k\rightarrow\infty$. Then, motivated by our earlier work

[17], we say that the sequence of mean values $J_{s,k}(X)$ $(k \in \mathbb{N})$ exhibits *quasi-diagonal behaviour* for the exponent s. It follows from [17, Theorem 1] that whenever $s \leqslant k^{3/2}(\log k)^{-1}$, quasi-diagonal behaviour holds for the mean value $J_{s,k}(X)$ in a particularly strong form. Indeed, subject to the latter condition on s, the bound (1.3) holds for the exponent $\delta_{s,k} = \exp(-Ak^3/s^2)$, for a certain positive constant A. In §9 we establish that the mean value $J_{s,k}(X)$ exhibits quasi-diagonal behaviour whenever $s = o(k^2)$.

Theorem 1.2. *Suppose that r, k and s are natural numbers with $k \geqslant 3$, $1 \leqslant r \leqslant \min\{k-2, \frac{1}{2}k+1\}$ and $s \leqslant r(k-r+2)$. Put*

$$\nu_{r,k} = \frac{r-1}{k-r}.$$

Then for each $\varepsilon > 0$, one has the estimate $J_{s,k}(X) \ll X^{s+\nu_{r,k}+\varepsilon}$.

In order to compare the strength of the estimate supplied by Theorem 1.2 with that of previous work, it is useful to consider the situation in which s and k are natural numbers with k large and $s \leqslant \frac{1}{4}k^2$, and to put $\lambda = s/k^2$. Then the work of Arkhipov and Karatsuba [2] shows that (1.3) holds with a permissible exponent $\delta_{s,k}$ satisfying $\delta_{s,k} \ll \lambda^{3/2}k^2$, Tyrina [11] obtains $\delta_{s,k} \ll \lambda^2 k^2$, whilst Theorem 1.2 yields the significantly stronger bound $\delta_{s,k} \ll \lambda$. Notice also that by taking $r = 1$ in Theorem 1.2, one recovers the estimate $J_{k+1,k}(X) \ll X^{k+1+\varepsilon}$ obtained in a slightly sharper form in Hua [8, Lemma 5.4], and sharpened further by Vaughan and Wooley [14]. Finally, by putting $r = [(k+1)/2]$ in Theorem 1.2, one obtains an attractive estimate simple to state.

Corollary 1.3. *Suppose that s and k are natural numbers with $k \geqslant 4$ and $s \leqslant \frac{1}{4}k^2 + k$. Then for each $\varepsilon > 0$, one has $J_{s,k}(X) \ll X^{s+1+\varepsilon}$.*

The estimate supplied by this corollary comes very close indeed to establishing the conjectured estimate (1.2) in the interval $1 \leqslant s \leqslant \frac{1}{4}k^2 + k$. If one were to establish an analogue of Corollary 1.3 in the longer interval $1 \leqslant s \leqslant \frac{1}{2}k(k+1)$, then the full conjecture (1.2) would essentially follow. In a sense, therefore, Corollary 1.3 comes half way to proving the main conjecture in this subject. When $s \geqslant k^2 - 1$, on the other hand, Theorem 1.1 establishes the conjectured bound (1.2). If one were to establish an analogue of Theorem 1.1 for $s \geqslant \frac{1}{2}k(k+1)$, this would again prove the main conjecture. Thus one comes half way to proving the main conjecture in two different directions.

The conclusion of Theorem 1.1 delivers essentially optimal estimates for $J_{s,k}(X)$ when $s \geqslant k^2 - 1$. In §8 we consider the behaviour of $J_{s,k}(X)$ when s is somewhat smaller than $k^2 - 1$. In this context, it is useful to define the exponent

$$\Delta_{t,k} = \tfrac{1}{2}t(t-1)\Big(\frac{k+1}{k-1}\Big). \tag{1.4}$$

Theorem 1.4. *Suppose that s, t and k are natural numbers with $k \geqslant 3$, $1 \leqslant t \leqslant k-1$ and $s \geqslant (k-t)(k+1)$. Then for each $\varepsilon > 0$, one has*

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

The exponent $\Delta_{t,k}$ in the upper bound presented in Theorem 1.4 converges quadratically to zero as t decreases to zero, representing a substantial improvement over the bounds made available by means of linear interpolation via Hölder's inequality. Notice that Theorem 1.1 follows from Theorem 1.4 by simply setting $t=1$.

We turn next to applications of our methods in the context of Waring's problem. When s and k are natural numbers, let $R_{s,k}(n)$ denote the number of representations of the natural number n as the sum of s kth powers of positive integers. A formal application of the circle method suggests that for $k\geqslant 3$ and $s\geqslant k+1$, one should have

$$R_{s,k}(n)=\frac{\Gamma(1+1/k)^s}{\Gamma(s/k)}\mathfrak{S}_{s,k}(n)n^{s/k-1}+o(n^{s/k-1}), \tag{1.5}$$

where

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

Subject to suitable congruence conditions, one has $1\ll \mathfrak{S}_{s,k}(n)\ll n^\varepsilon$, so that the conjectured relation (1.5) represents an honest asymptotic formula. Let $\widetilde{G}(k)$ denote the least integer t with the property that, for all $s\geqslant t$, and all sufficiently large natural numbers n, one has the asymptotic formula (1.5). By incorporating the estimates supplied by Theorems 1.1 and 1.4 into our recent work concerning the asymptotic formula in Waring's problem [21], in §10 we derive the upper bounds for $\widetilde{G}(k)$ contained in the following theorem. We make use here of the notation defined in (1.4).

Theorem 1.5. *Let k be a natural number with $k\geqslant 3$. Then one has*

$$\widetilde{G}(k)\leqslant 2k^2-2k+1-\max_{\substack{0\leqslant r\leqslant k-2\\ 2^r\leqslant k^2-k-1}}\left\lceil \frac{2(k-1)(r+1)-2^{r+1}}{k-r}\right\rceil,$$

and also

$$\widetilde{G}(k)\leqslant 2k^2-1-\max_{\substack{1\leqslant m\leqslant k\\ 2(t-1)(k+1)+m(m-1)<2k^2-2}}\max_{1\leqslant t\leqslant k-1}\left\lceil \frac{2(k+1)(t-1)-m(m-1)}{1+\Delta_{t,k}/m}\right\rceil.$$

Two consequences of Theorem 1.5 deserve to be recorded.

Corollary 1.6. *When k is a large natural number, one has*

$$\widetilde{G}(k)\leqslant 2k^2-k^{4/3}+O(k).$$

This conclusion sharpens slightly the bound $\widetilde{G}(k)\leqslant 2k^2-2\,[(\log k)/(\log 2)]$ established recently in [21, Corollary 1.2].

Corollary 1.7. *When k is a natural number with $k\geqslant 6$, one has*

$$\widetilde{G}(k)\leqslant 2k^2-2k-\theta_k,$$

where

$$\theta_k = \begin{cases} 8, & \textit{when } k = 6, \\ 9, & \textit{when } 7 \leqslant k \leqslant 13, \\ 10, & \textit{when } 14 \leqslant k \leqslant 19, \\ 12, & \textit{when } k \geqslant 20. \end{cases}$$

In particular, one has

$$\widetilde{G}(6) \leqslant 52,\ \widetilde{G}(7) \leqslant 75,\ \widetilde{G}(8) \leqslant 103,\ \widetilde{G}(9) \leqslant 135, \ldots,\ \widetilde{G}(20) \leqslant 748.$$

For comparison, in [21, Corollary 1.2] we showed that

$$\widetilde{G}(7) \leqslant 86, \widetilde{G}(8) \leqslant 117,\ \widetilde{G}(9) \leqslant 151, \ldots,\ \widetilde{G}(20) \leqslant 789.$$

Work preceding the introduction of efficient congruencing delivered substantially weaker conclusions. Thus, for smaller values of k, by using a refinement of an earlier method of Heath-Brown [7], it was shown by Boklan [3] that

$$\widetilde{G}(6) \leqslant 56,\ \widetilde{G}(7) \leqslant 112,\ \widetilde{G}(8) \leqslant 224.$$

For large values of k, meanwhile, one had the work of Ford [6]. Together with refinements for intermediate values of k due to Parsell [9] and Boklan and Wooley [4], this delivered the bounds

$$\widetilde{G}(9) \leqslant 365, \ldots, \widetilde{G}(20) \leqslant 2534, \quad \text{and} \quad \widetilde{G}(k) \leqslant k^2(\log k + \log\log k + O(1)).$$

We note that the methods underlying the proof of Theorem 1.5 fail by ε to deliver the bound $\widetilde{G}(5) \leqslant 32$ established by Vaughan [12]. Thus, our methods come within a whisker of achieving useful conclusions even for $k = 5$.

We establish Theorems 1.1, 1.2 and 1.4 by means of the *efficient congruencing* method introduced in our earlier work [20]. A sketch of the method is provided in [20, §2], and the reader may find this a helpful guide when it comes to understanding the basic plan of attack in this paper. It is a notable feature of this earlier work that, when successful for a given choice of s, the method yields a bound of the shape $J_{s,k}(X) \ll X^{2s-\frac{1}{2}k(k+1)+\varepsilon}$, within a factor X^ε of the sharpest bound conjectured to hold. In this paper we adapt the efficient congruencing method so as to obtain weaker bounds of the shape $J_{s,k}(X) \ll X^{2s-\kappa(s,k)+\varepsilon}$, wherein $\kappa(s,k) < \frac{1}{2}k(k+1)$. Although this advance may seem to provide only modest additional flexibility, it is neither trivial nor inconsequential. Further differences will be encountered from [20] in the handling of auxiliary congruences, and in particular linear congruence information is more efficiently handled implicitly within the main congruencing process.

We organise this paper as follows. In §2 we invest in some preliminary manoeuvres and introduce notation that facilitates what follows. Estimates for auxiliary congruences are established in §3, and in §4 we perform the conditioning of variables that permits non-singularity constraints to be imposed on the variables where needed. The efficient congruencing process is described in two stages. In §5 we perform the efficient congruencing step itself. Then, following discussion of an initial pre-congruencing step in §6, we advance in §7 to extract from the conclusions of §5 a formulation suitable for iterating the efficient congruencing process. We now come to the iterative relations, and these differ according to the variable regime of interest. In §§8 and 9 we establish, respectively, Theorems 1.1 and 1.4, and Theorem 1.2. Then in §10, we discuss the asymptotic formula in Waring's problem, proving Theorem 1.5 and its corollaries. Finally, in §11, we consider several further consequences of our new estimates. Here we highlight improvements in estimates of Weyl type, the distribution of polynomials modulo 1, Tarry's problem, and an estimate of Croot and Hart related to the sum-product theorem.

2. Preliminaries and infrastructure

Our objective in this section is to introduce such notation and preliminary estimates as are needed to describe the infrastructure of the repeated efficient congruencing process. In what follows, the letter k denotes a fixed integer exceeding 2, the letter s will be a positive integer, and ε denotes a sufficiently small positive number. The basic parameter occurring in our asymptotic estimates is X, a large real number depending at most on k, s and ε, unless otherwise indicated. In an effort to simplify our exposition, we adopt the following convention concerning the number ε. Whenever ε appears in a statement, either implicitly or explicitly, we assert that the statement holds for each $\varepsilon > 0$. Note that the "value" of ε may consequently change from statement to statement. We are relatively cavalier concerning the use of vector notation. In particular, we may write $\mathbf{z} \equiv \mathbf{w} \pmod{p}$ to denote that $z_i \equiv w_i \pmod{p}$ $(1 \leqslant i \leqslant t)$, or even $\mathbf{z} \equiv \xi \pmod{p}$ to denote that $z_i \equiv \xi \pmod{p}$ $(1 \leqslant i \leqslant t)$. Finally, throughout §§2–9, we consider the integer k to be fixed, and we therefore abbreviate $J_{s,k}(X)$ to $J_s(X)$, and likewise $f_k(\boldsymbol{\alpha}; X)$ to $f(\boldsymbol{\alpha}; X)$, without further comment.

Our attention is focused on the mean value $J_s(X)$ where, for the moment, we think of s as being an arbitrary natural number. We refer to the exponent λ_s as *permissible* when, for each positive number ε, and for any real number X sufficiently large in terms of s, k and ε, one has $J_s(X) \ll X^{\lambda_s+\varepsilon}$. Define λ_s^* to be the infimum of the set of exponents λ_s permissible for s and k. In view of the conjectured upper bound (1.2) and the corresponding lower bound (1.1), we expect that for each natural number s, one should have

$$\lambda_s^* = \max\{s, 2s - \tfrac{1}{2}k(k+1)\}.$$

In our earlier work [20], we sought to establish that $\lambda_s^* = 2s - \frac{1}{2}k(k+1)$ with s as small as possible, and indeed we established such for $s \geqslant k(k+1)$. In present circumstances we are less ambitious, though we ultimately prove more. With this in mind, we take $\kappa_s = \kappa(s,k)$ to be a positive parameter to be chosen in due course, but satisfying $\kappa_s \leqslant \max\{s, \frac{1}{2}k(k+1)\}$. In addition, we define $\eta_s = \eta_s(\kappa_s, k)$ by putting $\eta_s = \lambda_s^* - 2s + \kappa_s$. Thus, whenever X is sufficiently large in terms of s, k and ε, one has

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

where

$$\lambda_s^* = 2s - \kappa_s + \eta_s. \tag{2.2}$$

Rather than investigate the sequence of exponents λ_s^* directly, it is more convenient instead to fix a natural number r with

$$1 \leqslant r \leqslant k-1, \tag{2.3}$$

and then seek to bound λ_{s+r}^*. By choosing κ_{s+r} carefully in terms of s, we are able to apply the efficient congruencing process to show that η_{s+r} may be taken to be an arbitrarily small positive number, and thereby we demonstrate that in fact $\lambda_{s+r}^* \leqslant 2s+2r-\kappa_{s+r}$. We determine κ_{s+r} in terms of s and k by means of the parameter r as follows. Fix natural numbers s and s_0 with $s \geqslant s_0$, and write

$$\rho = k-r+1. \tag{2.4}$$

When it comes to proving Theorem 1.2 we take

$$s_0 = r\rho \quad \text{and} \quad \kappa_{s+r} = s_0 + r - \frac{r-1}{k-r}, \tag{2.5}$$

and for the proof of Theorem 1.4 we take

$$s_0 = rk \quad \text{and} \quad \kappa_{s+r} = \left(rk - \tfrac{1}{2}r(r+1)\right)\left(\frac{k+1}{k-1}\right). \tag{2.6}$$

Our goal is to show that $\lambda_{s+r}^* \leqslant 2(s+r)-\kappa_{s+r}$, and so we suppose by way of contradiction that in fact

$$\lambda_{s+r}^* = 2(s+r) - \kappa_{s+r} + \eta_{s+r},$$

with $\eta_{s+r} > 0$.

Let δ be a small positive number to be chosen shortly. In view of the infimal definition of λ_{s+r}^*, there exists a sequence of natural numbers $(X_n)_{n=1}^\infty$, tending to infinity, with the property that

$$J_{s+r}(X_n) > X_n^{\lambda_{s+r}^*-\delta} \quad (n \in \mathbb{N}). \tag{2.7}$$

Provided that X_n is sufficiently large, it follows from (2.1) that for $X_n^{\delta^2} < Y \leqslant X_n$, one has the corresponding upper bound

$$J_{s+r}(Y) < Y^{\lambda_{s+r}^*+\delta}. \tag{2.8}$$

Notice that since $s \geqslant s_0$, the trivial inequality $|f(\boldsymbol{\alpha};X)| \leqslant X$ yields the upper bound

$$J_{s+r}(X) \leqslant X^{2(s-s_0)} \oint |f(\boldsymbol{\alpha};X)|^{2s_0+2r}\,\mathrm{d}\boldsymbol{\alpha} = X^{2(s-s_0)} J_{s_0+r}(X).$$

Consequently, one has $\eta_{s+r} \leqslant \eta_{s_0+r}$, and so we are at liberty to restrict attention to the special case $s = s_0$. Since s_0 is a multiple of r, we consider a fixed natural number u with $u \geqslant s_0/r$, and put $s = ru$. We keep in play the general case $s \geqslant s_0$ until the final stages of our argument, the better to illuminate the underlying ideas. Finally, we take N to be a natural number sufficiently large in terms of s, k and r. In our proofs of Theorems 1.2 and 1.4 we put

$$\theta = N^{-1/2}(r/s)^{N+2} \tag{2.9}$$

and fix δ to be a positive number with $\delta < (Ns)^{-3N}$, so that δ is small compared to θ. We now take a fixed element $X = X_n$ of the sequence (X_n), which we may assume to be sufficiently large in terms of s, k, r, N and δ, and put $M=X^\theta$. In particular, we have $X^\delta<M^{1/N}$.

Let p be a fixed prime number with $M<p\leqslant 2M$ to be chosen in due course. That such a prime exists is a consequence of the Prime Number Theorem. When c and ξ are non-negative integers, and $\boldsymbol{\alpha}\in[0,1)^k$, define

$$\mathfrak{f}_c(\boldsymbol{\alpha};\xi)=\sum_{\substack{1\leqslant x\leqslant X\\ x\equiv \xi \pmod{p^c}}} e(\psi(x;\boldsymbol{\alpha})), \tag{2.10}$$

where

$$\psi(x;\boldsymbol{\alpha})=\alpha_1x+\alpha_2x^2+\ldots+\alpha_kx^k.$$

As in [20], we must consider well-conditioned tuples of integers belonging to distinct congruence classes modulo a suitable power of p, though now we must proceed in greater generality. Denote by $\Xi_c^r(\xi)$ the set of r-tuples $(\xi_1,\ldots,\xi_r)$, with

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

and satisfying the property that $\xi_i\equiv\xi_j\pmod{p^{c+1}}$ for no i and j with $1\leqslant i<j\leqslant r$. In addition, write $\Sigma_r=\{1,-1\}^r$, and consider an element $\boldsymbol{\sigma}$ of Σ_r. We then define

$$\mathfrak{F}_c^{\boldsymbol{\sigma}}(\boldsymbol{\alpha};\xi)=\sum_{\boldsymbol{\xi}\in\Xi_c^r(\xi)}\prod_{i=1}^r \mathfrak{f}_{c+1}(\sigma_i\boldsymbol{\alpha};\xi_i). \tag{2.11}$$

Notice that we have suppressed mention of the parameter r in our notation for the exponential sum $\mathfrak{F}_c^{\boldsymbol{\sigma}}(\boldsymbol{\alpha};\xi)$, based on the premise that any possible confusion should be easily avoided.

Two mixed mean values are important within our arguments. First, when a and b are positive integers and $\boldsymbol{\sigma}\in\Sigma_r$, we define

$$I_{a,b}^{\boldsymbol{\sigma}}(X;\xi,\eta)=\oint |\mathfrak{F}_a^{\boldsymbol{\sigma}}(\boldsymbol{\alpha};\xi)^2\mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{2s}|\,\mathrm{d}\boldsymbol{\alpha} \tag{2.12}$$

and

$$K_{a,b}^{\boldsymbol{\sigma},\boldsymbol{\tau}}(X;\xi,\eta)=\oint |\mathfrak{F}_a^{\boldsymbol{\sigma}}(\boldsymbol{\alpha};\xi)^2\mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)^{2u}|\,\mathrm{d}\boldsymbol{\alpha}. \tag{2.13}$$

It is convenient then to put

$$I_{a,b}(X)=\max_{1\leqslant\xi\leqslant p^a}\ \max_{\substack{1\leqslant\eta\leqslant p^b\\ \eta\not\equiv\xi \pmod{p}}}\ \max_{\boldsymbol{\sigma}\in\Sigma_r} I_{a,b}^{\boldsymbol{\sigma}}(X;\xi,\eta) \tag{2.14}$$

and

$$K_{a,b}(X)=\max_{1\leqslant\xi\leqslant p^a}\ \max_{\substack{1\leqslant\eta\leqslant p^b\\ \eta\not\equiv\xi \pmod{p}}}\ \max_{\boldsymbol{\sigma},\boldsymbol{\tau}\in\Sigma_r} K_{a,b}^{\boldsymbol{\sigma},\boldsymbol{\tau}}(X;\xi,\eta). \tag{2.15}$$

The implicit dependence of these mean values on our choice of p will ultimately be rendered irrelevant, since we fix p in the pre-congruencing step described in §6, following the proof of Lemma 6.1. We defer the definition of $K_{0,b}(X)$ to §6, since there are technical complications better avoided at this stage.

As in [20], our arguments are simplified by making transparent the relationship between mean values and their anticipated magnitudes. In this context, we define $[[J_{s+r}(X)]]$ by means of the relation

$$J_{s+r}(X)=X^{2s+2r-\kappa_{s+r}}[[J_{s+r}(X)]], \tag{2.16}$$

and when $0\leqslant a<b$, we define $[[K_{a,b}(X)]]$ by means of the relation

$$K_{a,b}(X)=(X/M^b)^{2s}(X/M^a)^{2r-\kappa_{s+r}}[[K_{a,b}(X)]]. \tag{2.17}$$

The lower bound (2.7) may now be written

$$[[J_{s+r}(X)]]>X^{\eta_{s+r}-\delta}. \tag{2.18}$$

We finish this section by recalling an estimate from [20] that encapsulates the translation-dilation invariance of the Diophantine system underlying the mean value $J_s(X)$.

Lemma 2.1. *Suppose that c is a non-negative integer with $c\theta\leqslant 1$. Then for each natural number t, one has*

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

Proof. This is [20, Lemma 3.1]. □

3. Auxiliary systems of congruences

Following the pattern established in our initial work [20] concerning efficient congruencing, we begin the main thrust of our analysis with a discussion of the congruences that play a critical role in what follows. Two basic arrangements of the congruencing idea are required, and these we handle in separate lemmata. We prepare the ground first with some notation.

Recall that r is an integer with $1\leqslant r\leqslant k-1$. When a and b are integers with $1\leqslant a<b$, and $\boldsymbol{\sigma}\in\Sigma_r$, we denote by $\mathcal{B}_{a,b}^{\boldsymbol{\sigma},r}(\mathbf{m};\xi,\eta)$ the set of solutions of the system of congruences

$$\sum_{i=1}^r\sigma_i(z_i-\eta)^j\equiv m_j\ (\text{mod }p^{jb})\quad(1\leqslant j\leqslant k), \tag{3.1}$$

with $1\leqslant\mathbf{z}\leqslant p^{kb}$ and $\mathbf{z}\equiv\boldsymbol{\xi}\ (\text{mod }p^{a+1})$ for some $\boldsymbol{\xi}\in\Xi_a^r(\xi)$. We define an equivalence relation $\mathcal{R}(\lambda)$ on integral r-tuples by declaring the r-tuples $\mathbf{x}$ and $\mathbf{y}$ to be $\mathcal{R}(\lambda)$-equivalent when $\mathbf{x}\equiv\mathbf{y}\ (\text{mod }p^\lambda)$. We then write $\mathcal{C}_{a,b}^{\boldsymbol{\sigma},r,h}(\mathbf{m};\xi,\eta)$ for the set of $\mathcal{R}(hb)$-equivalence classes of $\mathcal{B}_{a,b}^{\boldsymbol{\sigma},r}(\mathbf{m};\xi,\eta)$, and we define $B_{a,b}^{r,h}(p)$ by putting

$$B_{a,b}^{r,h}(p)=\max_{1\leqslant\xi\leqslant p^a}\max_{\substack{1\leqslant\eta\leqslant p^b\\ \eta\not\equiv\xi\ (\text{mod }p)}}\max_{\boldsymbol{\sigma}\in\Sigma_r}\max_{1\leqslant\mathbf{m}\leqslant p^{kb}}\text{card}(\mathcal{C}_{a,b}^{\boldsymbol{\sigma},r,h}(\mathbf{m};\xi,\eta)). \tag{3.2}$$

On considering representatives of the $\mathcal{R}(hb)$-equivalence classes of the set $\mathcal{B}_{a,b}^{\boldsymbol{\sigma},r}(\mathbf{m};\xi,\eta)$, of course, we may interpret $\mathcal{C}_{a,b}^{\boldsymbol{\sigma},r,h}(\mathbf{m};\xi,\eta)$ via the relation

$$\mathcal{C}_{a,b}^{\boldsymbol{\sigma},r,h}(\mathbf{m};\xi,\eta)=\{\mathbf{x}\ (\text{mod }p^{hb}):\mathbf{x}\in\mathcal{B}_{a,b}^{\boldsymbol{\sigma},r}(\mathbf{m};\xi,\eta)\}.$$

When $a=0$ we modify these definitions, so that $\mathcal{B}_{0,b}^{\boldsymbol{\sigma},r}(\mathbf{m};\xi,\eta)$ denotes the set of solutions of the system of congruences (3.1) with $1\leqslant \mathbf{z}\leqslant p^{kb}$ and $\mathbf{z}\equiv \boldsymbol{\xi} \pmod{p}$ for some $\boldsymbol{\xi}\in \Xi_0^r(\xi)$, and for which in addition one has $\mathbf{z}\not\equiv \eta \pmod{p}$. As in the previous case, we write $\mathcal{C}_{0,b}^{\boldsymbol{\sigma},r,h}(\mathbf{m};\xi,\eta)$ for the set of $\mathcal{R}(hb)$-equivalence classes of $\mathcal{B}_{0,b}^{\boldsymbol{\sigma},r}(\mathbf{m};\xi,\eta)$, but we define $B_{0,b}^{r,h}(p)$ by putting

$$B_{0,b}^{r,h}(p)=\max_{1\leqslant \eta\leqslant p^b}\max_{\boldsymbol{\sigma}\in\Sigma_r}\max_{1\leqslant \mathbf{m}\leqslant p^{kb}}\text{card}(\mathcal{C}_{0,b}^{\boldsymbol{\sigma},r,h}(\mathbf{m};0,\eta)). \tag{3.3}$$

We note that the choice of ξ in this situation with $a=0$ is irrelevant, since one has $\xi\equiv 0 \pmod{p^a}$ for all integers ξ. However, it is notationally convenient to preserve the similarity with the corresponding notation relevant to the situation with $a\geqslant 1$.

We aim to estimate $B_{a,b}^{r,h}(p)$ by exploiting the underlying non-singularity of the solution set via Hensel's lemma. A suitable version of the latter lifting process is implicitly contained within the following lemma.

Lemma 3.1. *Let $f_1,\ldots,f_d$ be polynomials in $\mathbb{Z}[x_1,\ldots,x_d]$ with respective degrees $k_1,\ldots,k_d$, and write*

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

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

$$f_j(x_1,\ldots,x_d)\equiv 0 \pmod{\varpi^l}\quad (1\leqslant j\leqslant d),$$

with $1\leqslant x_i\leqslant \varpi^l$ $(1\leqslant i\leqslant d)$ and $(J(\mathbf{f};\mathbf{x}),\varpi)=1$. Then $\mathcal{N}(\mathbf{f};\varpi^l)\leqslant k_1\cdots k_d$.

Proof. This is [19, Theorem 1]. □

We prepare a second auxiliary lemma in order to facilitate discussion of a certain argument involving elimination of terms amongst systems of polynomials. In this context, we adopt the convention that when l and m are natural numbers with $l>m$, then the binomial coefficient $\binom{m}{l}$ is zero.

Lemma 3.2. *Let α and β be natural numbers. Then there exist integers c_l $(\alpha\leqslant l\leqslant \alpha+\beta)$ and d_m $(\beta\leqslant m\leqslant \alpha+\beta)$, depending at most on α and β, and with $d_\beta\neq 0$, for which one has the polynomial identity*

$$c_\alpha+\sum_{l=1}^{\beta}c_{\alpha+l}(x+1)^{\alpha+l}=\sum_{m=\beta}^{\alpha+\beta}d_mx^m. \tag{3.4}$$

Proof. Consider the system of equations

$$\sum_{l=1}^{\beta}\binom{\alpha+l}{m}y_{\alpha+l}=\mu_m\quad (1\leqslant m\leqslant \beta), \tag{3.5}$$

in which μ_m is 0 when $1\leqslant m<\beta$, and 1 when $m=\beta$. By comparing coefficients of powers of x on left and right hand sides of (3.4), we see that the conclusion of the lemma follows provided that the system of linear equations

(3.5) admits a rational solution $\mathbf{y}$. Indeed, given such a solution, on taking d_β to be the least common multiple of the denominators of $y_{\alpha+l}$ $(1 \leqslant l \leqslant \beta)$, one finds that there exist integers c_α and d_m $(\beta < m \leqslant \alpha+\beta)$ for which the identity (3.4) holds with $c_{\alpha+l} = d_\beta y_{\alpha+l}$ $(1 \leqslant l \leqslant \beta)$.

We now demonstrate that the system (3.5) does indeed possess a rational solution. When $1 \leqslant m \leqslant \beta$, write

$$\psi_m(t) = t(t-1)\dots(t-m+1).$$

Then on multiplying the equations indexed by m in (3.5) through by $m!$, one finds that this system is equivalent to

$$\sum_{l=1}^{\beta} \psi_m(\alpha+l) y_{\alpha+l} = \beta! \mu_m \quad (1 \leqslant m \leqslant \beta).$$

Hence, on taking linear combinations of these equations, one discerns that (3.5) is in turn equivalent to the system of equations

$$\sum_{l=1}^{\beta} (\alpha+l)^m y_{\alpha+l} = \beta! \mu_m \quad (1 \leqslant m \leqslant \beta). \tag{3.6}$$

The matrix of coefficients of this system has determinant equal to the Vandermonde determinant

$$\det\left((\alpha+l)^m\right)_{1 \leqslant l,m \leqslant \beta} = \prod_{1 \leqslant l < m \leqslant \beta} ((\alpha+l)-(\alpha+m)) \neq 0,$$

and hence is invertible. We therefore deduce by means of Cramer's rule that the system (3.6) possesses a rational solution depending only on its coefficients, thus depending only on α and β. The same is consequently true of the equivalent system (3.5). In view of the discussion of the first paragraph, this suffices to complete the proof of the lemma. □

Our first bound for $B_{a,b}^{r,h}(p)$ addresses the scenario in which $r < k$, but $h = k$. In a sense, this situation is one in which we discard the $k-r$ congruences of smallest modulus p^{jb} $(1 \leqslant j \leqslant k-r)$ but nonetheless aim to lift solutions to the maximum modulus p^{kb}. This lemma must be prepared in two variants, one for the case $a \geqslant 1$ and a second for $a = 0$. Before announcing the lemma and its proof, we emphasise that throughout §§3-9, we assume r to be constrained by (2.3), and define ρ by means of (2.4).

Lemma 3.3. *Suppose that a and b are integers with $1 \leqslant a < b$. Then*

$$B_{a,b}^{r,k}(p) \leqslant k! p^{\frac{1}{2}r(r-1)(a+b)}.$$

Proof. Consider fixed integers a and b with $1 \leqslant a < b$, a fixed r-tuple $\boldsymbol{\sigma} \in \Sigma_r$, and fixed integers ξ and η with $1 \leqslant \xi \leqslant p^a$, $1 \leqslant \eta \leqslant p^b$ and $\eta \not\equiv \xi \pmod{p}$. We denote by $\mathcal{D}_1(\mathbf{n})$ the set of $\mathcal{R}(kb)$-equivalence classes of solutions of the system of congruences

$$\sum_{i=1}^{r} \sigma_i (z_i - \eta)^j \equiv n_j \pmod{p^{kb}} \quad (\rho \leqslant j \leqslant k), \tag{3.7}$$

with $1 \leqslant \mathbf{z} \leqslant p^{kb}$ and $\mathbf{z} \equiv \boldsymbol{\xi} \pmod{p^{a+1}}$ for some $\boldsymbol{\xi} \in \Xi_a^r(\xi)$. Given a fixed integral r-tuple $\mathbf{m}$, the number of r-tuples $\mathbf{n}$ with $1 \leqslant \mathbf{n} \leqslant p^{kb}$ for which

$$n_j \equiv m_j \pmod{p^{jb}} \quad (k-r+1 \leqslant j \leqslant k)$$

is equal to

$$\prod_{j=k-r+1}^{k} p^{(k-j)b} = (p^b)^{\frac{1}{2}r(r-1)}.$$

Consequently, it follows from (3.1) that

$$\begin{aligned}\operatorname{card}(\mathcal{C}_{a,b}^{\boldsymbol{\sigma},r,k}(\mathbf{m};\xi,\eta)) &\leqslant \sum_{\substack{1\leqslant n_\rho\leqslant p^{kb}\\ n_\rho\equiv m_\rho \pmod{p^{\rho b}}}} \dots \sum_{\substack{1\leqslant n_k\leqslant p^{kb}\\ n_k\equiv m_k \pmod{p^{kb}}}} \operatorname{card}(\mathcal{D}_1(\mathbf{n}))\\ &\leqslant (p^b)^{\frac{1}{2}r(r-1)} \max_{1\leqslant \mathbf{n}\leqslant p^{kb}} \operatorname{card}(\mathcal{D}_1(\mathbf{n})). \qquad (3.8)\end{aligned}$$

We next rewrite each variable z_i in the shape $z_i = p^a y_i + \xi$. In view of the hypothesis that $\mathbf{z} \equiv \boldsymbol{\xi} \pmod{p^{a+1}}$ for some $\boldsymbol{\xi} \in \Xi_a(\xi)$, the r-tuple $\mathbf{y}$ necessarily satisfies the property that

$$y_i \not\equiv y_m \pmod{p} \quad (1 \leqslant i < m \leqslant r). \tag{3.9}$$

Write $\zeta = \xi - \eta$, and note that the constraint $\eta \not\equiv \xi \pmod{p}$ ensures that $p \nmid \zeta$. It follows that there exists a multiplicative inverse of ζ modulo p^{kb}, and we denote this by ζ^{-1}. Then we deduce from (3.7) that $\operatorname{card}(\mathcal{D}_1(\mathbf{n}))$ is bounded above by the number of $\mathcal{R}(kb-a)$-equivalence classes of solutions of the system of congruences

$$\sum_{i=1}^{r} \sigma_i (p^a y_i \zeta^{-1} + 1)^j \equiv n_j (\zeta^{-1})^j \pmod{p^{kb}} \quad (\rho \leqslant j \leqslant k), \tag{3.10}$$

with $1 \leqslant \mathbf{y} \leqslant p^{kb-a}$ satisfying (3.9). Let $\mathbf{y} = \mathbf{w}$ be any solution of the system (3.10), if indeed such a solution exists. Then we find that all other solutions $\mathbf{y}$ satisfy the system of congruences

$$\sum_{i=1}^{r} \sigma_i \left((p^a y_i \zeta^{-1} + 1)^j - (p^a w_i \zeta^{-1} + 1)^j \right) \equiv 0 \pmod{p^{kb}} \quad (\rho \leqslant j \leqslant k). \tag{3.11}$$

It is at this point that we make use of Lemma 3.2. Consider an index j with $\rho \leqslant j \leqslant k$, and apply the latter lemma with $\alpha = \rho - 1$ and $\beta = j - \rho + 1$. We deduce that there exist integers c_{jl} $(\rho - 1 \leqslant l \leqslant j)$ and d_{jm} $(j-\rho+1 \leqslant m \leqslant j)$, depending at most on j and k, and with $d_{j,j-\rho+1} \neq 0$, for which one has the polynomial identity

$$c_{j,\rho-1} + \sum_{l=\rho}^{j} c_{jl}(x+1)^l = \sum_{m=j-\rho+1}^{j} d_{jm} x^m. \tag{3.12}$$

Since we may assume p to be sufficiently large in terms of $d_{j,j-\rho+1}$, moreover, there is no loss of generality in supposing that $p \nmid d_{j,j-\rho+1}$. Then by multiplying the equation (3.12) through by the multiplicative inverse of $d_{j,j-\rho+1}$ modulo p^{kb}, we see that there is no loss in supposing that $d_{j,j-\rho+1} \equiv 1 \pmod{p^{kb}}$. By taking suitable linear combinations of the congruences comprising (3.11), we thus infer that any solution of this system satisfies

$$(\zeta^{-1}p^a)^{j-\rho+1}\sum_{i=1}^r \sigma_i(\psi_j(y_i)-\psi_j(w_i))\equiv 0 \pmod{p^{kb}}\quad (\rho\leqslant j\leqslant k),$$

in which we have written

$$\psi_j(z)=z^{j-\rho+1}+\sum_{m=j-\rho+2}^{j} d_{jm}(\zeta^{-1}p^a)^{m-j+\rho-1}z^m. \tag{3.13}$$

Note here, in particular, that

$$\psi_j(z)\equiv z^{j-\rho+1} \pmod{p}. \tag{3.14}$$

Denote by $\mathcal{D}_2(\mathbf{u})$ the set of $\mathcal{R}(kb-a)$-equivalence classes of solutions of the system of congruences

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

with $1\leqslant \mathbf{y}\leqslant p^{kb-a}$ satisfying (3.9). Then we have shown thus far that

$$\text{card}(\mathcal{D}_1(\mathbf{n}))\leqslant \max_{1\leqslant \mathbf{u}\leqslant p^{kb}} \text{card}(\mathcal{D}_2(\mathbf{u})). \tag{3.15}$$

Let $\mathcal{D}_3(\mathbf{v})$ denote the set of $\mathcal{R}(kb-a)$-equivalence classes of solutions of the system

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

with $1\leqslant \mathbf{y}\leqslant p^{kb-a}$ satisfying (3.9). Then

$$\begin{aligned}\text{card}(\mathcal{D}_2(\mathbf{u}))&\leqslant \sum_{\substack{1\leqslant v_\rho\leqslant p^{kb-a}\\ v_\rho\equiv u_\rho \pmod{p^{kb-a}}}}\dots\sum_{\substack{1\leqslant v_k\leqslant p^{kb-a}\\ v_k\equiv u_k \pmod{p^{kb-ra}}}}\text{card}(\mathcal{D}_3(\mathbf{v}))\\ &\leqslant (p^a)^{\frac12 r(r-1)}\max_{1\leqslant \mathbf{v}\leqslant p^{kb-a}}\text{card}(\mathcal{D}_3(\mathbf{v})).\end{aligned} \tag{3.16}$$

Define the determinant

$$J(\boldsymbol{\psi};\mathbf{x})=\det\left(\sigma_i\psi'_{\rho+l-1}(x_i)\right)_{1\leqslant i,l\leqslant r}. \tag{3.17}$$

We claim that when $y_i\equiv y_m \pmod{p}$ for no i and m with $1\leqslant i<m\leqslant r$, then $(J(\boldsymbol{\psi};\mathbf{y}),p)=1$. Temporarily assuming the validity of this claim, we deduce from Lemma 3.1 that $\text{card}(\mathcal{D}_3(\mathbf{v}))\leqslant \rho(\rho+1)\cdots k\leqslant k!$. In view of the definition (3.2), the conclusion of the lemma follows at once from (3.8), (3.15) and (3.16).

In order to confirm the validity of our claim concerning the Jacobian determinant, we begin by observing that (3.14) implies that

$$\sigma_i\psi'_{\rho+l-1}(y_i)\equiv \sigma_i l y_i^{l-1} \pmod{p}.$$

Since we have supposed p to be large compared to k, we find that $p|J(\boldsymbol{\psi};\mathbf{y})$ if and only if

$$\det(y_i^{l-1})_{1\leqslant i,l\leqslant r}\equiv 0 \pmod{p}.$$

But by hypothesis we have $y_i\equiv y_m \pmod{p}$ for no i and m with $1\leqslant i<m\leqslant r$, and so it follows that

$$\det(y_i^{l-1})_{1\leqslant i,l\leqslant r}=\prod_{1\leqslant i<m\leqslant r}(y_i-y_m)\not\equiv 0 \pmod{p}.$$

We are therefore forced to conclude that $p\nmid J(\boldsymbol{\psi};\mathbf{y})$, thereby confirming the validity of our earlier claim, and completing the proof of the lemma. □

A variant of Lemma 3.3 supplies an analogue applicable in the case $a=0$.

Lemma 3.4. *Suppose that b is an integer with $b\geqslant 1$. Then*

$$B_{0,b}^{r,k}(p)\leqslant k!p^{\frac{1}{2}r(r-1)b}.$$

Proof. Consider a fixed integer b with $b\geqslant 1$, a fixed r-tuple $\boldsymbol{\sigma}\in\Sigma_r$, and a fixed integer η with $1\leqslant\eta\leqslant p^b$. We denote by $\mathcal{D}_1(\mathbf{n};\eta)$ the set of $\mathcal{R}(kb)$-equivalence classes of solutions of the system of congruences (3.7) with $1\leqslant\mathbf{z}\leqslant p^{kb}$ and $\mathbf{z}\equiv\boldsymbol{\xi} \pmod{p}$ for some $\boldsymbol{\xi}\in\Xi_0^r(0)$, and for which in addition $\mathbf{z}\not\equiv\eta \pmod{p}$. Then as in the opening paragraph of the proof of Lemma 3.3, it follows from (3.1) that

$$\text{card}(\mathcal{C}_{0,b}^{\boldsymbol{\sigma},r,k}(\mathbf{m};0,\eta))\leqslant(p^b)^{\frac{1}{2}r(r-1)}\max_{1\leqslant\mathbf{n}\leqslant p^{kb}}\text{card}(\mathcal{D}_1(\mathbf{n};\eta)).\tag{3.18}$$

But $\mathcal{D}_1(\mathbf{n};\eta)=\mathcal{D}_1(\mathbf{n};0)$, and $\mathcal{D}_1(\mathbf{n};0)$ counts the solutions of the system of congruences

$$\sum_{i=1}^r\sigma_iy_i^j\equiv n_j \pmod{p^{kb}}\quad(\rho\leqslant j\leqslant k),$$

with $1\leqslant\mathbf{y}\leqslant p^{kb}$ satisfying (3.9), and in addition $p\nmid y_i$ $(1\leqslant i\leqslant r)$. Write

$$J(\mathbf{y})=\det\left((\rho+j-1)\sigma_iy_i^{\rho+j-2}\right)_{1\leqslant i,j\leqslant r}.\tag{3.19}$$

Then since p is large compared to k, we find that $p|J(\mathbf{y})$ if and only if

$$(y_1\dots y_r)^{\rho-1}\det\left(y_i^{j-1}\right)_{1\leqslant i,j\leqslant r}\equiv 0 \pmod{p}.$$

But by hypothesis we have $(y_1\dots y_r,p)=1$ and $y_i\equiv y_j \pmod{p}$ for no i and j with $1\leqslant i<j\leqslant r$, and so it follows that

$$(y_1\dots y_r)^{\rho-1}\det\left(y_i^{j-1}\right)_{1\leqslant i,j\leqslant r}=(y_1\dots y_r)^{\rho-1}\prod_{1\leqslant i<j\leqslant r}(y_i-y_j)\not\equiv 0 \pmod{p}.$$

We therefore deduce from Lemma 3.1 that $\mathcal{D}_1(\mathbf{n};0)\leqslant\rho(\rho+1)\dots k\leqslant k!$. In view of (3.3), the conclusion of the lemma therefore follows from (3.18). □

Our second bound for $B_{a,b}^{r,h}(p)$ addresses the scenario in which $h=k-r+1$ and $r<k$. This situation amounts to one in which we aim to lift solutions to an intermediate modulus $p^{\rho b}$, and discard any congruences of modulus smaller than $p^{\rho b}$. Again, we provide two variants of this lemma, one with $a\geqslant 1$ and a second with $a=0$.

Lemma 3.5. *Suppose that a and b are natural numbers with $b \geqslant (r-1)a$. Then $B_{a,b}^{r,\rho}(p) \leqslant k! p^{(r-1)a}$.*

Proof. Consider fixed natural numbers a and b with $b \geqslant (r-1)a$, a fixed r-tuple $\boldsymbol{\sigma} \in \Sigma_r$, and fixed integers ξ and η with $1 \leqslant \xi \leqslant p^a$, $1 \leqslant \eta \leqslant p^b$ and $\eta \not\equiv \xi \pmod{p}$. In addition, define the integer μ_j for $\rho \leqslant j \leqslant k$ by putting

$$\mu_j = \begin{cases} 0, & \text{when } \rho+1 \leqslant j \leqslant k,\\ r-1, & \text{when } j = \rho. \end{cases}$$

We denote by $\mathcal{D}_1(\mathbf{n})$ the set of $\mathcal{R}(\rho b)$-equivalence classes of solutions of the system of congruences

$$\sum_{i=1}^r \sigma_i (z_i - \eta)^j \equiv n_j \pmod{p^{jb+\mu_j a}} \quad (\rho \leqslant j \leqslant k), \tag{3.20}$$

with $1 \leqslant \mathbf{z} \leqslant p^{\rho b}$ and $\mathbf{z} \equiv \boldsymbol{\xi} \pmod{p^{a+1}}$ for some $\boldsymbol{\xi} \in \Xi_a^r(\xi)$. Then it follows from (3.1) that

$$\begin{aligned} \operatorname{card}(\mathcal{C}_{a,b}^{\boldsymbol{\sigma},r,\rho}(\mathbf{m};\xi,\eta)) &\leqslant \sum_{\substack{1 \leqslant n \leqslant p^{\rho b+(r-1)a}\\ n \equiv m_\rho \ (\mathrm{mod}\ p^{\rho b})}} \operatorname{card}(\mathcal{D}_1(n, m_{\rho+1}, \ldots, m_k))\\ &\leqslant p^{(r-1)a} \max_{1 \leqslant \mathbf{n} \leqslant p^{kb}} \operatorname{card}(\mathcal{D}_1(\mathbf{n})). \end{aligned} \tag{3.21}$$

Following the pattern of the proof of Lemma 3.3, we next rewrite each variable z_i in the shape $z_i = p^a y_i + \xi$. The hypothesis that $\mathbf{z} \equiv \boldsymbol{\xi} \pmod{p^{a+1}}$ for some $\boldsymbol{\xi} \in \Xi_a^r(\xi)$ again implies that the r-tuple $\mathbf{y}$ satisfies (3.9). Let $\zeta = \xi - \eta$ and write ζ^{-1} for the multiplicative inverse of ζ modulo p^{kb}. Then we deduce from (3.20) that $\operatorname{card}(\mathcal{D}_1(\mathbf{n}))$ is bounded above by the number of $\mathcal{R}(\rho b - a)$-equivalence classes of solutions of the system of congruences

$$\sum_{i=1}^r \sigma_i (p^a y_i \zeta^{-1} + 1)^j \equiv n_j (\zeta^{-1})^j \pmod{p^{\rho b+(r-1)a}} \quad (\rho \leqslant j \leqslant k), \tag{3.22}$$

with $1 \leqslant \mathbf{y} \leqslant p^{\rho b - a}$ satisfying (3.9). Here, we have made use of the fact that since $b \geqslant (r-1)a$, then for $j \geqslant \rho+1$ the validity of a congruence modulo p^{jb} implies that of the corresponding congruence modulo $p^{\rho b+(r-1)a}$.

Let $\mathbf{y} = \mathbf{w}$ be any solution of the system (3.22), if such a solution exists. Then we find that all other solutions $\mathbf{y}$ satisfy the system of congruences

$$\sum_{i=1}^r \sigma_i \left((p^a y_i \zeta^{-1} + 1)^j - (p^a w_i \zeta^{-1} + 1)^j \right) \equiv 0 \pmod{p^{\rho b+(r-1)a}} \quad (\rho \leqslant j \leqslant k). \tag{3.23}$$

Recall the definition (3.13) of the polynomials $\psi_j(z)$. Then by taking linear combinations of these congruences, we find as in the proof of Lemma 3.3 that there exist integers d_{jm} $(j-\rho+2 \leqslant m \leqslant j)$, for $\rho \leqslant j \leqslant k$, with the property that any solution of (3.23) satisfies the system of congruences

$$(\zeta^{-1}p^a)^{j-\rho+1}\sum_{i=1}^r \sigma_i\left(\psi_j(y_i)-\psi_j(w_i)\right)\equiv 0 \pmod{p^{\rho b+(r-1)a}}\quad (\rho\leqslant j\leqslant k). \tag{3.24}$$

Denote by $\mathcal{D}_2(\mathbf{u})$ the set of $\mathcal{R}(\rho b-a)$-equivalence classes of solutions of the system of congruences

$$\sum_{i=1}^r \sigma_i\psi_j(y_i)\equiv u_j \pmod{p^{\rho b-a}}\quad (\rho\leqslant j\leqslant k), \tag{3.25}$$

with $1\leqslant \mathbf{y}\leqslant p^{\rho b-a}$ satisfying (3.9). Note that when $\rho\leqslant j\leqslant k$, one has

$$j-\rho+1\leqslant k-(k-r+1)+1=r.$$

Then it follows from (3.24) that

$$\text{card}(\mathcal{D}_1(\mathbf{n}))\leqslant \max_{1\leqslant \mathbf{u}\leqslant p^{kb}}\text{card}(\mathcal{D}_2(\mathbf{u})). \tag{3.26}$$

With the Jacobian determinant $J(\boldsymbol{\psi};\mathbf{x})$ defined as in (3.17), we find as in the proof of Lemma 3.3 that the solutions $\mathbf{y}$ of (3.25) counted by $\mathcal{D}_2(\mathbf{u})$ satisfy $(J(\boldsymbol{\psi};\mathbf{y}),p)=1$. We therefore deduce from Lemma 3.1 that $\text{card}(\mathcal{D}_2(\mathbf{u}))\leqslant \rho(\rho+1)\dots k\leqslant k!$. In view of (3.2), the conclusion of the lemma now follows from (3.21) and (3.26). □

Again, a variant of Lemma 3.5 supplies an analogue applicable in the special case $a=0$.

Lemma 3.6. *Suppose that b is an integer with $b\geqslant 1$. Then $B^{r,\rho}_{0,b}(p)\leqslant k!$.*

Proof. Consider a fixed integer b with $b\geqslant 1$, a fixed r-tuple $\boldsymbol{\sigma}\in\Sigma_r$, and a fixed integer η with $1\leqslant\eta\leqslant p^b$. We denote by $\mathcal{D}_1(\mathbf{n};\eta)$ the set of $\mathcal{R}(\rho b)$-equivalence classes of solutions of the system of congruences

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

with $1\leqslant \mathbf{z}\leqslant p^{\rho b}$ and $\mathbf{z}\equiv\boldsymbol{\xi}\pmod p$ for some $\boldsymbol{\xi}\in\Xi_0^r(0)$, and for which in addition $\mathbf{z}\not\equiv\eta\pmod p$. Then it follows from (3.1) that

$$\text{card}(\mathcal{C}^{\boldsymbol{\sigma},r,\rho}_{0,b}(\mathbf{m};0,\eta))\leqslant \max_{1\leqslant\mathbf{n}\leqslant p^{\rho b}}\text{card}(\mathcal{D}_1(\mathbf{n};\eta)). \tag{3.27}$$

Recall the definition of the Jacobian determinant $J(\mathbf{y})$ from (3.19). Then following the argument concluding the proof of Lemma 3.4, one discerns that $\mathcal{D}_1(\mathbf{n};\eta)=\mathcal{D}_1(\mathbf{n};0)$, and that $\mathcal{D}_1(\mathbf{n};0)$ counts the solutions of the system of congruences

$$\sum_{i=1}^r \sigma_i y_i^j\equiv n_j \pmod{p^{\rho b}}\quad (\rho\leqslant j\leqslant k),$$

with $1\leqslant\mathbf{y}\leqslant p^{\rho b}$ satisfying $p\nmid J(\mathbf{y})$. By wielding Lemma 3.1, we therefore deduce that $\mathcal{D}_1(\mathbf{n};0)\leqslant\rho(\rho+1)\dots k\leqslant k!$. In view of (3.3), the conclusion of the lemma therefore follows from (3.27). □

4. The conditioning process

As in the analogous treatment of [20, §5], the mean value $I_{a,b}^{\boldsymbol{\sigma}}(X;\xi,\eta)$ is not, by itself, suitable for use in a repeated efficient congruencing iteration. In this section we show how, without serious loss, one may replace the factor $\mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{2s}$ occurring in (2.12) by the conditioned factor $\mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)^{2u}$ in (2.13). Our argument follows very closely the proof of [20, Lemma 5.1], and so we may be concise by analogy at several points in our discussion.

Lemma 4.1. *Let a and b be integers with $b > a \geqslant 1$. Then one has*

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

Proof. Consider fixed integers ξ and η with $1 \leqslant \xi \leqslant p^a$ and $1 \leqslant \eta \leqslant p^b$ with $\eta \not\equiv \xi \pmod p$, and an r-tuple $\boldsymbol{\sigma} \in \Sigma_r$. Then on considering the underlying Diophantine system, it follows from (2.12) that $I_{a,b}^{\boldsymbol{\sigma}}(X;\xi,\eta)$ counts the number of integral solutions of the system

$$\sum_{i=1}^r \sigma_i (x_i^j - y_i^j) = \sum_{l=1}^s (v_l^j - w_l^j) \quad (1 \leqslant j \leqslant k), \tag{4.1}$$

with

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

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

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

Let T_1 denote the number of integral solutions $\mathbf{x}$, $\mathbf{y}$, $\mathbf{v}$, $\mathbf{w}$ of the system (4.1), counted by $I_{a,b}^{\boldsymbol{\sigma}}(X;\xi,\eta)$, in which the $2s$ integers $v_1,\ldots,v_s$ and $w_1,\ldots,w_s$ together lie in at most $r-1$ distinct residue classes modulo p^{b+1}, and let T_2 denote the corresponding number of solutions in which these integers together occupy at least r distinct residue classes modulo p^{b+1}. Then

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

The argument of the proof of [20, Lemma 5.1] leading to equation (5.2) of that paper shows, mutatis mutandis, that

$$\begin{aligned} T_1 &\ll \sum_{\substack{1 \leqslant \eta_1,\ldots,\eta_{r-1} \leqslant p^{b+1} \\ \boldsymbol{\eta} \equiv \eta \pmod{p^b}}} \sum_{i=1}^r \oint |\mathfrak{F}_a^{\boldsymbol{\sigma}}(\boldsymbol{\alpha};\xi)^2 \mathfrak{f}_{b+1}(\boldsymbol{\alpha};\eta_i)^{2s}| \, \mathrm{d}\boldsymbol{\alpha} \\ &\ll p^{r-1} \max_{\substack{1 \leqslant \eta_0 \leqslant p^{b+1} \\ \eta_0 \not\equiv \xi \pmod p}} I_{a,b+1}^{\boldsymbol{\sigma}}(X;\xi,\eta_0). \end{aligned}$$

On the other hand, the argument of the proof of [20, Lemma 5.1] leading to equation (5.3) of that paper shows, mutatis mutandis, that for some $\boldsymbol{\tau} \in \Sigma_r$ one has

$$\begin{aligned} T_2 &\ll \left(\oint |\mathfrak{F}_a^{\boldsymbol{\sigma}}(\boldsymbol{\alpha};\xi)^2 \mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)^{2u}| \, \mathrm{d}\boldsymbol{\alpha} \right)^{1/(2u)} \left(\oint |\mathfrak{F}_a^{\boldsymbol{\sigma}}(\boldsymbol{\alpha};\xi)^2 \mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{2s}| \, \mathrm{d}\boldsymbol{\alpha} \right)^{1-1/(2u)} \\ &\ll \left(K_{a,b}^{\boldsymbol{\sigma},\boldsymbol{\tau}}(X;\xi,\eta) \right)^{1/(2u)} \left(I_{a,b}^{\boldsymbol{\sigma}}(X;\xi,\eta) \right)^{1-1/(2u)}. \end{aligned}$$

Thus we deduce from (2.14) and (2.15) that

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

The conclusion of the lemma follows immediately. □

We next obtain an estimate that enables us to truncate the conditioning process. Here we recall that the exponent κ_{s+r} is a positive number, with

$$\kappa_{s+r}\leqslant \max\{s+r,\tfrac{1}{2}k(k+1)\},$$

which measures the strength of the permissible exponent λ_s^* by means of the relation (2.2). We have in mind the choices for κ_{s+r} presented in equations (2.5) and (2.6). Finally, it is convenient to write κ for κ_{s+r}, since confusion is easily avoided.

Lemma 4.2. *Let a, b and H be positive integers with*

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

Then provided that $s\geqslant 3r$, one has

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

Proof. On considering the underlying Diophantine equations, we find from (2.12) that when $1\leqslant \xi\leqslant p^a$, $1\leqslant \eta\leqslant p^{b+H}$ and $\boldsymbol{\sigma}\in\Sigma_r$, one has

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

Applying Hölder's inequality together with Lemma 2.1, therefore, we obtain

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

We thus deduce from (2.2) and (2.8) that

$$\begin{aligned}I_{a,b+H}(X)&\ll \bigl((X/M^a)^{r/(s+r)}(X/M^{b+H})^{s/(s+r)}\bigr)^{2s+2r-\kappa+\eta_{s+r}+\delta}\\ &\ll X^\delta (X/M^a)^{2r-\kappa+\eta_{s+r}}(X/M^b)^{2s}\Upsilon,\end{aligned}$$

where

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

We may suppose that $s+r\geqslant \kappa$, $H\geqslant 2(b-a)$ and $s\geqslant 3r$, and hence

$$\begin{aligned}rH+(b-a+H)\kappa s/(s+r)-2sH&\leqslant rH+\tfrac{3}{2}\kappa sH/(s+r)-2sH\\ &\leqslant \tfrac{1}{2}(2r-s)H\leqslant -\tfrac{1}{2}rH.\end{aligned}$$

Consequently, one has

$$M^{H(r-1)}\Upsilon\leqslant M^{-(r+2)H/2},$$

whence

$$M^{H(r-1)}I_{a,b+H}(X)\ll M^{-(r+2)H/2}X^\delta (X/M^a)^{2r-\kappa+\eta_{s+r}}(X/M^b)^{2s},$$

and the conclusion of the lemma follows. □

The repeated application of Lemma 4.1 in combination with Lemma 4.2 yields the conditioning lemma underpinning the efficient congruencing process.

Lemma 4.3. *Let a and b be integers with $1 \leqslant a < b$, and put $H = 2(b-a)$. Suppose that $b + H \leqslant \theta^{-1}$ and $s \geqslant 3r$. Then there exists an integer h with $0 \leqslant h < H$ having the property that*

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

Proof. Repeated application of Lemma 4.1 shows that whenever a and b are positive integers with $b > a \geqslant 1$, and $H = 2(b-a)$, then

$$I_{a,b}(X) \ll \sum_{h=0}^{H-1} M^{h(r-1)} K_{a,b+h}(X) + M^{H(r-1)} I_{a,b+H}(X). \tag{4.2}$$

The desired conclusion therefore follows on applying Lemma 4.2 to estimate the second term on the right hand side of (4.2). □

5. The efficient congruencing step, I

Our goal in this section is to convert latent congruence information within the mean value $K_{a,b}(X)$ into a form useful in subsequent iterations, and this we achieve using the work of §3. The two basic approaches of §3 yield two different manifestations of the efficient congruencing step, and these we examine in separate lemmata.

Lemma 5.1. *Suppose that a and b are integers with $1 \leqslant a < b \leqslant \theta^{-1}$. Then one has*

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

Proof. Consider fixed integers ξ and η with $1 \leqslant \xi \leqslant p^a$, $1 \leqslant \eta \leqslant p^b$ and $\eta \not\equiv \xi \pmod{p}$, and r-tuples $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \Sigma_r$. Then by orthogonality, the mean value $K_{a,b}^{\boldsymbol{\sigma};\boldsymbol{\tau}}(X;\xi,\eta)$ defined in (2.13) counts the number of integral solutions of the system

$$\sum_{i=1}^{r} \sigma_i (x_i^j - y_i^j) = \sum_{l=1}^{u} \sum_{m=1}^{r} \tau_m (v_{lm}^j - w_{lm}^j) \quad (1 \leqslant j \leqslant k), \tag{5.1}$$

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

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

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

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

As in the argument of the proof of [20, Lemma 6.1], an application of the Binomial Theorem shows that these solutions satisfy the system of congruences

$$\sum_{i=1}^{r} \sigma_i (x_i - \eta)^j \equiv \sum_{i=1}^{r} \sigma_i (y_i - \eta)^j \pmod{p^{jb}} \quad (1 \leqslant j \leqslant k). \tag{5.2}$$

We now make use of the work of §3, writing

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

Then on considering the underlying Diophantine system, we see from (5.1) and (5.2) that

$$K^{\boldsymbol{\sigma},\boldsymbol{\tau}}_{a,b}(X;\xi,\eta)=\sum_{m_1=1}^{p^b}\dots\sum_{m_k=1}^{p^{kb}}\oint|\mathfrak{G}^{\boldsymbol{\sigma}}_{a,b}(\boldsymbol{\alpha};\xi,\eta;\mathbf{m})^2\mathfrak{F}^{\boldsymbol{\tau}}_b(\boldsymbol{\alpha};\eta)^{2u}|\,\mathrm{d}\boldsymbol{\alpha}.\tag{5.3}$$

An application of Cauchy's inequality leads via Lemma 3.3 to the bound

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

whence

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

As in the argument of the proof of [20, Lemma 6.1] leading to equation (6.7) of the latter paper, from here an application of Hölder's inequality yields the upper bound

$$\begin{aligned}&K^{\boldsymbol{\sigma},\boldsymbol{\tau}}_{a,b}(X;\xi,\eta)\\&\quad\ll M^{\frac{1}{2}r(r-1)(a+b)}(M^{kb-a})^r\max_{\substack{1\leqslant\zeta\leqslant p^{kb}\\ \zeta\equiv\xi\ (\mathrm{mod}\ p^a)}}\oint|\mathfrak{f}_{kb}(\boldsymbol{\alpha};\zeta)^{2r}\mathfrak{F}^{\boldsymbol{\tau}}_b(\boldsymbol{\alpha};\eta)^{2u}|\,\mathrm{d}\boldsymbol{\alpha}.\end{aligned}\tag{5.4}$$

Next we apply Hölder's inequality to the integral on the right hand side of (5.4) to obtain

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

where

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

and

$$U_2=I^{\boldsymbol{\tau}}_{b,kb}(X;\eta,\zeta).$$

Notice here that since $\eta\not\equiv\xi\pmod{p}$ and $\zeta\equiv\xi\pmod{p^a}$ with $a\geqslant 1$, we have $\zeta\not\equiv\eta\pmod{p}$. In this way we deduce from Lemma 2.1 that

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

and the conclusion of the lemma follows from (5.4). □

A variant of the argument employed to establish Lemma 5.1 makes use of Lemma 3.5 in place of Lemma 3.3.

Lemma 5.2. *Suppose that a and b are integers with $1 \leqslant a < b \leqslant \theta^{-1}$ and $b \geqslant (r-1)a$. Then one has*

$$K_{a,b}(X) \ll M^{(r-1)a}(M^{\rho b-a})^r (J_{s+r}(X/M^b))^{1-r/s}(I_{b,\rho b}(X))^{r/s}.$$

Proof. Initially we follow the argument of the proof of Lemma 5.1, identifying $K_{a,b}^{\boldsymbol{\sigma},\boldsymbol{\tau}}(X;\xi,\eta)$ with the number of integral solutions of the system (5.1) with its attendant conditions, and observing that the system of congruences (5.2) necessarily holds for each solution. We now write

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

and note as before that the relation (5.3) again holds. An application of Cauchy's inequality in this instance leads from Lemma 3.5 to the estimate

$$\begin{aligned}|\mathfrak{G}_{a,b}^{\boldsymbol{\sigma}}(\boldsymbol{\alpha};\xi,\eta;\mathbf{m})|^2 &\leqslant \operatorname{card}(\mathcal{C}_{a,b}^{\boldsymbol{\sigma},r,\rho}(\mathbf{m};\xi,\eta)) \sum_{\boldsymbol{\zeta}\in\mathcal{C}_{a,b}^{\boldsymbol{\sigma},r,\rho}(\mathbf{m};\xi,\eta)} \prod_{i=1}^r |\mathfrak{f}_{\rho b}(\boldsymbol{\alpha};\zeta_i)|^2\\ &\ll M^{(r-1)a} \sum_{\boldsymbol{\zeta}\in\mathcal{C}_{a,b}^{\boldsymbol{\sigma},r,\rho}(\mathbf{m};\xi,\eta)} \prod_{i=1}^r |\mathfrak{f}_{\rho b}(\boldsymbol{\alpha};\zeta_i)|^2,\end{aligned}$$

and hence

$$K_{a,b}^{\boldsymbol{\sigma},\boldsymbol{\tau}}(X;\xi,\eta) \ll M^{(r-1)a} \sum_{\substack{1\leqslant\boldsymbol{\zeta}\leqslant p^{\rho b}\\ \boldsymbol{\zeta}\equiv\xi \pmod{p^a}}} \oint \Bigl(\prod_{i=1}^r |\mathfrak{f}_{\rho b}(\boldsymbol{\alpha};\zeta_i)|^2\Bigr) |\mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)|^{2u}\,\mathrm{d}\boldsymbol{\alpha}.$$

From here, as in the argument leading to (5.4) above, one obtains

$$\begin{aligned}&K_{a,b}^{\boldsymbol{\sigma},\boldsymbol{\tau}}(X;\xi,\eta)\\ &\qquad\ll M^{(r-1)a}(M^{\rho b-a})^r \max_{\substack{1\leqslant\zeta\leqslant p^{\rho b}\\ \zeta\equiv\xi \pmod{p^a}}} \oint |\mathfrak{f}_{\rho b}(\boldsymbol{\alpha};\zeta)^{2r}\mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)^{2u}|\,\mathrm{d}\boldsymbol{\alpha}. \qquad (5.5)\end{aligned}$$

Applying Hölder's inequality as in the concluding paragraph of the proof of Lemma 5.1, we deduce that

$$\oint |\mathfrak{f}_{\rho b}(\boldsymbol{\alpha};\zeta)^{2r}\mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)^{2u}|\,\mathrm{d}\boldsymbol{\alpha} \ll (J_{s+r}(X/M^b))^{1-r/s}(I_{b,\rho b}(X))^{r/s},$$

and the conclusion of the lemma follows from (5.5). □

A crude but simple upper bound for $K_{a,b}(X)$ is useful in simplifying the argument to come.

Lemma 5.3. *Suppose that a and b are integers with $0 \leqslant a < b \leqslant \theta^{-1}$. Then*

$$[[K_{a,b}(X)]] \ll X^{\eta_{s+r}+\delta}(M^{b-a})^{\kappa}.$$

Proof. We adapt the argument of the proof of [20, Lemma 6.2]. Consider fixed integers ξ and η with $1 \leqslant \xi \leqslant p^a$ and $1 \leqslant \eta \leqslant p^b$, and r-tuples $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \Sigma_r$. Then from (2.13) it follows by orthogonality combined with Hölder's inequality that

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

Consequently, Lemma 2.1 delivers the bound

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

whence

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

This completes the proof of the lemma. □

6. The pre-congruencing step

In order to fix choices for ξ and η in §§3–5, one must first initiate the congruencing process. It is here that the choice for the prime number p is fixed once and for all. Before delving further into the details of this pre-congruencing step, we pause to introduce some additional notation. We amend the definition of the set $\Xi_c^r(\xi)$ from the discussion leading to (2.11) as follows. When $\mathrm{H} \subseteq \{1,\dots,p\}$, we denote by $\Xi(\mathrm{H})$ the set of r-tuples $(\xi_1,\dots,\xi_r)$ satisfying $1 \leqslant \boldsymbol{\xi} \leqslant p$ and in addition the property that one has neither $\xi_i \equiv \zeta \pmod{p}$ for any $\zeta \in \mathrm{H}$ $(1 \leqslant i \leqslant r)$, nor $\xi_i \equiv \xi_j \pmod{p}$ for any i and j with $1 \leqslant i < j \leqslant r$. Recalling (2.10), we next define the exponential sum $\mathfrak{F}(\boldsymbol{\alpha};\mathrm{H})$ by putting

$$\mathfrak{F}(\boldsymbol{\alpha};\mathrm{H}) = \sum_{\boldsymbol{\xi} \in \Xi(\mathrm{H})} \prod_{i=1}^r \mathfrak{f}_1(\boldsymbol{\alpha};\xi_i). \tag{6.1}$$

Also, when H is a subset of $\{1,\dots,p\}$ with cardinality $k-r$, we write

$$L(\boldsymbol{\alpha};\mathrm{H}) = \prod_{\zeta\in\mathrm{H}} \mathfrak{f}_1(\boldsymbol{\alpha};\zeta).$$

Finally, we write

$$\widetilde{I}_c(X;\eta) = \oint |\mathfrak{F}(\boldsymbol{\alpha};\{\eta\})^2 \mathfrak{f}_c(\boldsymbol{\alpha};\eta)^{2s}| \,\mathrm{d}\boldsymbol{\alpha}, \tag{6.2}$$

$$\widetilde{K}_c^{\boldsymbol{\tau}}(X;\eta) = \oint |\mathfrak{F}(\boldsymbol{\alpha};\{\eta\})^2 \mathfrak{F}_c^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)^{2u}| \,\mathrm{d}\boldsymbol{\alpha}, \tag{6.3}$$

$$\widetilde{K}_c(X) = \max_{1\leqslant\eta\leqslant p^c} \max_{\boldsymbol{\tau}\in\Sigma_r} K_c^{\boldsymbol{\tau}}(X;\eta). \tag{6.4}$$

Lemma 6.1. *Suppose that $s \geqslant \max\{k+1-r, 2r\}$ and $\kappa \leqslant s+r$. Then there exists a prime number p with $M < p \leqslant 2M$, and an integer $h \in \{0,1\}$, for which one has*

$$J_{s+r}(X) \ll M^{2s+h(r-1)} \widetilde{K}_{1+h}(X).$$

Proof. We adapt the argument of the proof of [20, Lemma 3.2]. The quantity $J_{s+r}(X)$ counts the number of integral solutions of the system

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

with $1 \leqslant \mathbf{x}, \mathbf{y} \leqslant X$. Let T_0 denote the number of such solutions in which $x_i = x_m$ for some i and m with $1 \leqslant i < m \leqslant k$, and let T_1 denote the corresponding number of solutions with $x_i = x_m$ for no i and m with $1 \leqslant i < m \leqslant k$. Then $J_{s+r}(X) = T_0 + T_1$.

Write $\Xi^*(p)$ for the set of k-tuples $\boldsymbol{\xi} = (\xi_1, \dots, \xi_k)$, with $1 \leqslant \boldsymbol{\xi} \leqslant p$, and satisfying the property that $\xi_i \equiv \xi_m \pmod{p}$ for no i and m with $1 \leqslant i < m \leqslant k$. In addition, define the exponential sum $\mathfrak{F}^*(\boldsymbol{\alpha})$ by putting

$$\mathfrak{F}^*(\boldsymbol{\alpha}) = \sum_{\boldsymbol{\xi} \in \Xi^*(p)} \prod_{i=1}^{k} \mathfrak{f}_1(\boldsymbol{\alpha}; \xi_i),$$

and write

$$I^*(X) = \oint |\mathfrak{F}^*(\boldsymbol{\alpha})^2 \mathfrak{f}_0(\boldsymbol{\alpha}; 0)^{2s+2r-2k}| \, \mathrm{d}\boldsymbol{\alpha}. \tag{6.5}$$

Then the argument of the proof of [20, Lemma 3.2] leading to equations (3.14) and (3.15) of the latter paper reveals that a prime number p exists, with $M < p \leqslant 2M$, for which

$$T_0 \ll (J_{s+r}(X))^{1-1/(2s+2r)} \quad \text{and} \quad T_1 \ll (I^*(X))^{1/2} (J_{s+r}(X))^{1/2}.$$

We thus infer that

$$J_{s+r}(X) \ll 1 + I^*(X) \ll I^*(X). \tag{6.6}$$

Next, splitting the summation in the definition (2.10) of $\mathfrak{f}_0(\boldsymbol{\alpha}; 0)$ into arithmetic progressions modulo p and applying Hölder's inequality, we obtain

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

It therefore follows from (6.5) that

$$I^*(X) \ll M^{2s+2r-2k} \max_{1 \leqslant \eta \leqslant p} T_2(\eta), \tag{6.7}$$

where

$$T_2(\eta) = \oint |\mathfrak{F}^*(\boldsymbol{\alpha})^2 \mathfrak{f}_1(\boldsymbol{\alpha}; \eta)^{2s+2r-2k}| \, \mathrm{d}\boldsymbol{\alpha}.$$

By orthogonality, the mean value $T_2(\eta)$ counts the integral solutions of the system

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

with

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

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

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

Consider a fixed choice of $\boldsymbol{\xi} \in \Xi^*(p)$. One has $\xi_l \equiv \eta \pmod{p}$ for at most one index l with $1 \leqslant l \leqslant k$. Since we suppose that $1 \leqslant r \leqslant k-1$, it follows that one may relabel indices in such a way that $(\xi_1, \ldots, \xi_r) \in \Xi_0^r(0)$ and $\xi_i \equiv \eta \pmod{p}$ for no index i with $1 \leqslant i \leqslant r$. One may do likewise with the variables $\mathbf{y}$. Notice that when $(\xi_1, \ldots, \xi_k) \in \Xi^*(p)$ and $(\xi_1, \ldots, \xi_r) \in \Xi_0^r(0)$, then necessarily $\xi_i \equiv \xi_j \pmod{p}$ for no indices i and j with $1 \leqslant i \leqslant r$ and $r+1 \leqslant j \leqslant k$. On considering the underlying Diophantine equations, therefore, we find that

$$T_2(\eta) \ll \oint \Biggl| \sum_{\substack{\mathrm{H} \subseteq \{1,\ldots,p\} \\ \mathrm{card}(\mathrm{H}) = k-r}} \mathfrak{F}(\boldsymbol{\alpha}; \mathrm{H} \cup \{\eta\}) L(\boldsymbol{\alpha}; \mathrm{H}) \Biggr|^2 |\mathfrak{f}_1(\boldsymbol{\alpha}; \eta)|^{2s+2r-2k} \, \mathrm{d}\boldsymbol{\alpha}.$$

An application of the elementary inequality

$$|z_1 \ldots z_n| \leqslant |z_1|^n + \ldots + |z_n|^n \tag{6.8}$$

reveals that

$$L(\boldsymbol{\alpha}; \mathrm{H}) \ll \sum_{\zeta \in \mathrm{H}} |\mathfrak{f}_1(\boldsymbol{\alpha}; \zeta)|^{k-r},$$

and thus we deduce via Cauchy's inequality that

$$T_2(\eta) \ll p^{k-r} \sum_{\substack{\mathrm{H} \subseteq \{1,\ldots,p\} \\ \mathrm{card}(\mathrm{H}) = k-r}} \sum_{\zeta \in \mathrm{H}} \oint |\mathfrak{F}(\boldsymbol{\alpha}; \mathrm{H} \cup \{\eta\}) \mathfrak{f}_1(\boldsymbol{\alpha}; \zeta)^{k-r} \mathfrak{f}_1(\boldsymbol{\alpha}; \eta)^{s+r-k}|^2 \, \mathrm{d}\boldsymbol{\alpha}.$$

A second application of (6.8) shows that

$$|\mathfrak{f}_1(\boldsymbol{\alpha}; \zeta)^{k-r} \mathfrak{f}_1(\boldsymbol{\alpha}; \eta)^{s+r-k}|^2 \ll |\mathfrak{f}_1(\boldsymbol{\alpha}; \zeta)|^{2s} + |\mathfrak{f}_1(\boldsymbol{\alpha}; \eta)|^{2s},$$

and hence we conclude that

$$T_2(\eta) \ll p^{2k-2r} \max_{\substack{\mathrm{H} \subseteq \{1,\ldots,p\} \\ \mathrm{card}(\mathrm{H}) = k-r}} \max_{\zeta \in \mathrm{H} \cup \{\eta\}} \oint |\mathfrak{F}(\boldsymbol{\alpha}; \mathrm{H} \cup \{\eta\})^2 \mathfrak{f}_1(\boldsymbol{\alpha}; \zeta)^{2s}| \, \mathrm{d}\boldsymbol{\alpha}.$$

Finally, a consideration of the underlying Diophantine system permits the last estimate to be simplified, so that on recalling (6.2) we arrive at the bound

$$\begin{aligned} T_2(\eta) &\ll p^{2k-2r} \max_{1 \leqslant \zeta \leqslant p} \oint |\mathfrak{F}(\boldsymbol{\alpha}; \{\zeta\})^2 \mathfrak{f}_1(\boldsymbol{\alpha}; \zeta)^{2s}| \, \mathrm{d}\boldsymbol{\alpha} \\ &\ll M^{2k-2r} \max_{1 \leqslant \zeta \leqslant p} \widetilde{I}_1(X; \zeta). \end{aligned}$$

Returning to (6.6) and (6.7), we may thus conclude that

$$J_{s+r}(X) \ll I^*(X) \ll M^{2s} \max_{1\leqslant \eta\leqslant p} \widetilde{I}_1(X;\eta). \tag{6.9}$$

The mean value $\widetilde{I}_1(X;\eta)$ counts the number of integral solutions of the system

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

with

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

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

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

Let T_3 denote the number of such solutions in which the $2s$ integers $v_1,\dots,v_s$ and $w_1,\dots,w_s$ together occupy at least r distinct residue classes modulo p^2, and let T_4 denote the corresponding number of solutions in which these integers together lie in at most $r-1$ distinct residue classes modulo p^2. Then we see that

$$\widetilde{I}_1(X;\eta) = T_3 + T_4. \tag{6.10}$$

The argument of the proof of [20, Lemma 5.1] leading to equation (5.3) of that paper shows, mutatis mutandis, that for some $\boldsymbol{\tau}\in\Sigma_r$, one has

$$T_3 \ll \Bigl(\oint |\mathfrak{F}(\boldsymbol{\alpha};\{\eta\})^2 \mathfrak{F}_1^{\boldsymbol{\tau}}(\boldsymbol{\alpha};\eta)^{2u}|\,\mathrm{d}\boldsymbol{\alpha}\Bigr)^{1/(2u)} \times \Bigl(\oint |\mathfrak{F}(\boldsymbol{\alpha};\{\eta\})^2 \mathfrak{f}_1(\boldsymbol{\alpha};\eta)^{2s}|\,\mathrm{d}\boldsymbol{\alpha}\Bigr)^{1-1/(2u)}.$$

Then on recalling (6.2) and (6.3), we deduce from (6.10) that

$$\widetilde{I}_1(X;\eta) \ll \Bigl(\widetilde{K}_1^{\boldsymbol{\tau}}(X;\eta)\Bigr)^{1/(2u)} \Bigl(\widetilde{I}_1(X;\eta)\Bigr)^{1-1/(2u)} + T_4,$$

whence

$$\widetilde{I}_1(X;\eta) \ll \widetilde{K}_1^{\boldsymbol{\tau}}(X;\eta) + T_4. \tag{6.11}$$

On the other hand, the argument of the proof of [20, Lemma 5.1] leading to equation (5.2) of that paper shows that

$$T_4 \ll \sum_{\substack{1\leqslant \eta_1,\dots,\eta_{r-1}\leqslant p^2\\ \boldsymbol{\eta}\equiv \eta \pmod{p}}} \sum_{i=1}^{r-1} \oint |\mathfrak{F}(\boldsymbol{\alpha};\{\eta\})^2 \mathfrak{f}_2(\boldsymbol{\alpha};\eta_i)^{2s}|\,\mathrm{d}\boldsymbol{\alpha}.$$

Such a conclusion may also be extracted from the argument of the proof of Lemma 4.1 above. A consideration of the underlying Diophantine system therefore shows that

$$T_4 \ll M^{r-1} \max_{1\leqslant \zeta\leqslant p^2} \oint |\mathfrak{F}(\boldsymbol{\alpha};\{\zeta\})^2 \mathfrak{f}_2(\boldsymbol{\alpha};\zeta)^{2s}|\,\mathrm{d}\boldsymbol{\alpha}.$$

In view of the relation (6.11), therefore, we deduce that

$$\widetilde{I}_1(X;\eta) \ll \widetilde{K}_1^{\boldsymbol{\tau}}(X;\eta) + M^{r-1} \max_{1\leqslant \zeta \leqslant p^2} \widetilde{I}_2(X;\zeta). \tag{6.12}$$

We may analyse the mean value $\widetilde{I}_2(X;\zeta)$ just as in our treatment of $\widetilde{I}_1(X;\eta)$ above, and thus we deduce that for some $\boldsymbol{\tau} \in \Sigma_r$, one has

$$\widetilde{I}_2(X;\zeta) \ll \widetilde{K}_2^{\boldsymbol{\tau}}(X;\zeta) + M^{r-1} \max_{1\leqslant \theta \leqslant p^3} \widetilde{I}_3(X;\theta).$$

On recalling (6.4), therefore, we find from (6.12) that

$$\widetilde{I}_1(X;\eta) \ll \widetilde{K}_1(X) + M^{r-1}\widetilde{K}_2(X) + M^{2r-2} \max_{1\leqslant \eta \leqslant p^3} \widetilde{I}_3(X;\eta).$$

Thus we deduce from (6.9) that there exists an integer $h \in \{0,1\}$ for which one has

$$J_{s+r}(X) \ll M^{2s+h(r-1)}\widetilde{K}_{1+h}(X) + M^{2s+2r-2} \max_{1\leqslant \eta \leqslant p^3} \widetilde{I}_3(X;\eta). \tag{6.13}$$

Next, on considering the underlying Diophantine system, an application of Hölder's inequality in combination with Lemma 2.1 confirms that

$$\begin{aligned}\widetilde{I}_3(X;\eta) &\ll \Big(\oint |\mathfrak{f}_0(\boldsymbol{\alpha};0)|^{2s+2r}\,\mathrm{d}\boldsymbol{\alpha}\Big)^{r/(s+r)} \Big(\max_{1\leqslant \eta \leqslant p^3} \oint |\mathfrak{f}_3(\boldsymbol{\alpha};\eta)|^{2s+2r}\,\mathrm{d}\boldsymbol{\alpha}\Big)^{s/(s+r)} \\ &\ll (X^{2s+2r-\kappa+\delta})^{r/(s+r)} \left((X/M^3)^{2s+2r-\kappa+\delta}\right)^{s/(s+r)}.\end{aligned}$$

Hence we have

$$M^{2s+2r-2}\widetilde{I}_3(X;\eta) \ll X^{2s+2r-\kappa+\delta}M^{\omega},$$

where

$$\omega = 2s+2r-2-\frac{3s}{s+r}(2s+2r-\kappa+\delta) \leqslant 2r-2-4s+3s\kappa/(s+r).$$

But by hypothesis, one has $\kappa \leqslant s+r$ and $s \geqslant 2r$, and thus

$$\omega \leqslant (2r-2-4s)+3s \leqslant -2.$$

Consequently, we derive the upper bound

$$M^{2s+2r-2}\widetilde{I}_3(X;\eta) \ll X^{2s+2r-\kappa-2\delta} \ll X^{-\delta}J_{s+r}(X).$$

On recalling (6.13), therefore, we see that there exists an integer $h \in \{0,1\}$ for which

$$J_{s+r}(X) \ll X^{-\delta}J_{s+r}(X) + M^{2s+h(r-1)}\widetilde{K}_{1+h}(X).$$

The conclusion of the lemma follows at once. □

We now fix the prime number p, once and for all, so that the upper bound for $J_{s+r}(X)$ claimed in the conclusion of Lemma 6.1 holds. In analysing the iterative process, we shall find it useful to have available versions of Lemmata 5.1 and 5.2 valid also when $a=0$. It is for this purpose that we prepared Lemma 6.1, as we now make transparent. In this context, we view the mean value $\widetilde{K}_c^{\boldsymbol{\tau}}(X;\eta)$ defined in (6.3) as a surrogate for $K_{0,c}^{\boldsymbol{1},\boldsymbol{\tau}}(X;0,\eta)$. Here, the implicit condition on variables avoiding the congruence class η modulo p, captured through the exponential sum $\mathfrak{F}(\boldsymbol{\alpha};\{\eta\})$ defined in (6.1), provides the correct analogue of the condition $\xi \not\equiv \eta \pmod{p}$. With this discussion in mind, we henceforth adopt the convention that when $b \in \{1,2\}$, one is to interpret the expression $K_{0,b}(X)$ as $\widetilde{K}_b(X)$.

Lemma 6.2. *Suppose that a and b are integers with $0 \leqslant a < b \leqslant \theta^{-1}$. Suppose further that $s \geqslant \max\{k+1-r, 2r\}$ and $\kappa \leqslant s+r$, and that when $a = 0$ one has $b = 1$ or 2. Then*

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

Proof. When $a \geqslant 1$ the conclusion asserted by the lemma is an immediate consequence of that supplied by Lemma 5.1. We therefore focus attention on the situation in which $a = 0$ and $b \in \{1,2\}$. Here, we find from (6.3) and (6.4) that

$$K_{0,b}(X) = \widetilde{K}_b(X) = \max_{1 \leqslant \eta \leqslant p^b} \max_{\boldsymbol{\tau} \in \Sigma_r} \oint |\mathfrak{F}(\boldsymbol{\alpha}; \{\eta\})^2 \mathfrak{F}_b^{\boldsymbol{\tau}}(\boldsymbol{\alpha}; \eta)^{2u}| \, \mathrm{d}\boldsymbol{\alpha}. \tag{6.14}$$

Imitating the argument of the proof of Lemma 5.1, and substituting the application of Lemma 3.4 for our earlier use of Lemma 3.3, with the discussion of the preamble to the present lemma in mind, we find that

$$K_{0,b}(X) \ll M^{\frac{1}{2}r(r-1)b}(M^{kb})^r(J_{s+r}(X/M^b))^{1-r/s}(I_{b,kb}(X))^{r/s},$$

and thus the desired conclusion does indeed hold when $a = 0$ and $b \in \{1,2\}$. □

Lemma 6.3. *Suppose that a and b are integers with $0 \leqslant a < b \leqslant \theta^{-1}$ and $b \geqslant (r-1)a$. Suppose further that $s \geqslant \max\{k+1-r, 2r\}$ and $\kappa \leqslant s+r$, and that when $a = 0$ one has $b = 1$ or 2. Then*

$$K_{a,b}(X) \ll M^{(r-1)a}(M^{\rho b-a})^r(J_{s+r}(X/M^b))^{1-r/s}(I_{b,\rho b}(X))^{r/s}.$$

Proof. In this instance, when $a \geqslant 1$ the conclusion asserted by the lemma follows from Lemma 5.2. We therefore focus again on the situation in which $a = 0$ and $b \in \{1,2\}$. We again find from (6.3) and (6.4) that the relation (6.14) holds. In present circumstances, by imitating the argument of the proof of Lemma 5.2, and substituting the application of Lemma 3.6 for our earlier use of Lemma 3.5, with the discussion of the preamble to Lemma 6.2 in mind, we find that

$$K_{0,b}(X) \ll (M^{\rho b})^r(J_{s+r}(X/M^b))^{1-r/s}(I_{b,\rho b}(X))^{r/s}.$$

This yields the desired conclusion when $a = 0$ and $b \in \{1,2\}$, and completes the proof of the lemma. □

7. The efficient congruencing step, II

By means of Lemmata 6.2 and 6.3, one is able to relate $K_{a,b}(X)$ either to $I_{b,kb}(X)$ or $I_{b,\rho b}(X)$, the purpose of the pre-congruencing step being to remove the constraint $a \geqslant 1$ imposed in §5 so as to permit a to be zero. In this section we complete the discussion of the efficient congruencing step by combining Lemma 4.3 first with Lemma 6.2, and then with Lemma 6.3, so as to obtain the basic iterative relations between $K_{a,b}(X)$ and $K_{b,kb+h}(X)$ in the first instance, and between $K_{a,b}(X)$ and $K_{b,\rho b+h}(X)$ in the second instance.

Lemma 7.1. *Define s_0 and κ as in (2.6), and put $s = s_0$. Suppose that a and b are integers with $0 \leqslant a < b \leqslant \frac{1}{3}(k\theta)^{-1}$, and put $H = 2(k-1)b$ and $g = b - ka$. Suppose further that when $a = 0$, one has $b = 1$ or 2. Then there exists an integer h, with $0 \leqslant h < H$, having the property that*

$$\begin{aligned}[][K_{a,b}(X)]] \ll & X^\delta \left(M^{sg-(2s-r+1)h} [[K_{b,kb+h}(X)]] \right)^{r/s} (X/M^b)^{\eta_{s+r}(1-r/s)} \\ & + M^{-rH/(3s)} (X/M^b)^{\eta_{s+r}}.\end{aligned}$$

Proof. We assume throughout that $k \geqslant 3$ and $1 \leqslant r \leqslant k-1$, so we may begin with the observation that $s = rk \geqslant \max\{3r, k-r+1\}$. Next, since $\frac{1}{2}r(r+1) \geqslant r$ for $r \geqslant 1$, we find from (2.6) that

$$\kappa = (rk - \tfrac{1}{2}r(r+1)) \left(\frac{k+1}{k-1} \right) \leqslant (rk - r) \left(\frac{k+1}{k-1} \right) = rk + r,$$

so that $\kappa \leqslant s + r$. On recalling (2.17), we may therefore apply Lemma 6.2 to deduce that

$$[[K_{a,b}(X)]] \ll (M^b)^{2s} (M^a)^{2r-\kappa} M^{\frac{1}{2}r(r-1)(b+a)} (M^{kb-a})^r T_1^{1-r/s} T_2^{r/s}, \tag{7.1}$$

where

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

But

$$T_1 \ll (M^{-b})^{2s+2r-\kappa} (X/M^b)^{\eta_{s+r}+\delta}. \tag{7.2}$$

Writing $H = 2(k-1)b$, we find that the hypotheses of the statement of the lemma guarantee that $kb + H = (3k-2)b < \theta^{-1}$. We therefore see from Lemma 4.3 that there exists an integer h with $0 \leqslant h < H$ such that

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

Fixing this value of h, we have

$$T_2 \ll (M^{-kb})^{2s} (M^{-b})^{2r-\kappa} \Omega, \tag{7.3}$$

where

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

On combining (7.1), (7.2) and (7.3), we conclude that

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

where

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

On recalling (2.6), a brief computation reveals that

$$\begin{aligned}\omega(a,b)&=(\kappa-rk+\tfrac{1}{2}r(r-1))b-(\kappa-\tfrac{1}{2}r(r+1))a\\&=\left(\frac{r(k-r)}{k-1}\right)(b-ka)=\left(\frac{r(k-r)}{k-1}\right)g.\end{aligned}$$

Consequently, on the one hand we have $\omega(a,b)\leqslant rg$, and on the other

$$\begin{aligned}\omega(a,b)s/r-(r+1)H/2&\leqslant\frac{s(k-r)b}{k-1}-(r+1)(k-1)b\\&\leqslant(kr-(r+1)(k-1))b=-(k-r-1)b\leqslant 0.\end{aligned}$$

Thus we infer that

$$\begin{aligned}[\![K_{a,b}(X)]\!]\ll{}&(M^{-H/2})^{r/s}X^\delta(X/M^b)^{\eta_{s+r}}\\&+X^\delta M^{rg-(2s-r+1)hr/s}(X/M^b)^{\eta_{s+r}(1-r/s)}[\![K_{b,kb+h}(X)]\!]^{r/s}.\end{aligned}$$

The conclusion of the lemma follows on observing that δ may be assumed small enough that $X^\delta\ll M^{rH/(6s)}$. □

Lemma 7.2. *Suppose that* $1\leqslant r\leqslant\min\{k-2,\frac{1}{2}k+1\}$*, define* s_0 *and* κ *as in (2.5), and put* $s=s_0$*. Suppose that* a *and* b *are integers with* $b\geqslant(r-1)a$ *and* $0\leqslant a<b\leqslant\frac{1}{3}(\rho\theta)^{-1}$*, and put* $H=2(\rho-1)b$ *and* $g=b-\rho a$*. Suppose further that when* $a=0$*, one has* $b=1$ *or* 2*. Then there exists an integer* h*, with* $0\leqslant h<H$*, having the property that*

$$\begin{aligned}[\![K_{a,b}(X)]\!]\ll{}&X^\delta\left(M^{sg-(2s-r+1)h}[\![K_{b,\rho b+h}(X)]\!]\right)^{r/s}(X/M^b)^{\eta_{s+r}(1-r/s)}\\&+M^{-rH/(3s)}(X/M^b)^{\eta_{s+r}}.\end{aligned}$$

Proof. We now assume that $k\geqslant 3$ and $1\leqslant r\leqslant\min\{k-2,\frac{1}{2}k+1\}$, so we have

$$s=r\rho=r(k+1-r)\geqslant\max\{3r,k-r+1\}.$$

Next, we see from (2.5) that

$$\kappa=s_0+r-\frac{r-1}{k-r}\leqslant s+r.$$

On recalling (2.17), we may therefore apply Lemma 6.3 to deduce that

$$[\![K_{a,b}(X)]\!]\ll(M^b)^{2s}(M^a)^{2r-\kappa}M^{(r-1)a}(M^{\rho b-a})^rT_1^{1-r/s}T_2^{r/s},\tag{7.4}$$

where in this instance

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

Writing $H=2(\rho-1)b$, the hypotheses of the statement of the lemma imply that $\rho b+H\leqslant(3\rho-2)b<\theta^{-1}$. Thus we deduce from Lemma 4.3 that there exists an integer h with $0\leqslant h<H$ such that

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

Fixing this value of h, we see that

$$T_2\ll(M^{-\rho b})^{2s}(M^{-b})^{2r-\kappa}\Omega,\tag{7.5}$$

where

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

On combining (7.4) and (7.5) with the estimate (7.2), still valid in the present setting, we reach the upper bound

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

where

$$\begin{aligned}\omega(a,b) = &\,2sb + (2r-\kappa)a + (r-1)a + r(\rho b - a)\\ &- (1-r/s)(2s+2r-\kappa)b - (2s\rho b + (2r-\kappa)b)r/s.\end{aligned}$$

We next recall (2.5), and hence deduce that

$$\omega(a,b) = (\kappa - r\rho)b + (2r-\kappa-1)a = \left(r - \frac{r-1}{\rho-1}\right)(b-\rho a).$$

Consequently, on the one hand we have $\omega(a,b) \leqslant rg$, and on the other

$$\begin{aligned}\omega(a,b)s/r - (r+1)H/2 &\leqslant \rho\Big(r - \frac{r-1}{\rho-1}\Big)b - (r+1)(\rho-1)b\\ &\leqslant (\rho r - r + 1)b - (\rho r + \rho - r - 1)b\\ &= -(k-r-1)b \leqslant 0.\end{aligned}$$

Thus we conclude that

$$\begin{aligned}[[K_{a,b}(X)]] \ll &\,(M^{-H/2})^{r/s}X^\delta (X/M^b)^{\eta_{s+r}}\\ &+ X^\delta M^{rg-(2s-r+1)hr/s}(X/M^b)^{\eta_{s+r}(1-r/s)}[[K_{b,\rho b+h}(X)]]^{r/s}.\end{aligned}$$

Just as in the conclusion of the proof of the previous lemma, our argument is completed by noting the estimate $X^\delta \ll M^{rH/(6s)}$. □

8. The iterative process, I: the basic estimate

Making use of Lemma 6.1, and then applying either Lemma 7.1 repeatedly, or else Lemma 7.2 repeatedly, we are able to bound $J_{s+r}(X)$ in terms of quantities of the shape $K_{c,d}(X)$, wherein c and d pass through an increasing sequence of integral values. Our goal in this section is to control this iterative process so as to establish Theorems 1.1 and 1.4. Although we model this treatment on the analogous analysis of [20, §7], there are complications in the details that generate some complexity.

Lemma 8.1. *Define s_0 and κ as in (2.6), and put $s = s_0$. Let a and b be integers with $0 \leqslant a < b \leqslant \frac{1}{3}(k\theta)^{-1}$ having the property that when $a = 0$, one has $b = 1$ or 2, and put $g = b - ka$. Suppose in addition that there exist non-negative numbers ψ, c and γ, with $c \leqslant 3(s/r)^N$, for which*

$$X^{\eta_{s+r}(1+\psi\theta)} \ll X^{c\delta}M^{-\gamma}[[K_{a,b}(X)]]. \tag{8.1}$$

Then, for some non-negative integer h with $h \leqslant 2(k-1)b$, one has

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

where

$$\psi' = (s/r)\psi + (s/r-1)b, \quad c' = (s/r)(c+1),$$
$$a' = b, \quad b' = kb+h, \quad \gamma' = (s/r)\gamma + (2s-r+1)h - sg.$$

Proof. Since we may suppose that $c \leqslant 3(s/r)^N$ and $\delta < (Ns)^{-3N}$, we have $c\delta < \theta/(6s)$, and hence $X^{c\delta} < M^{1/(6s)}$. In addition, one has $M^{1/(6s)} > X^\delta$. We therefore deduce from Lemma 7.1 that there exists an integer h with $0 \leqslant h < 2(k-1)b$ with the property that

$$[[K_{a,b}(X)]] \ll X^\delta (X/M^b)^{(1-r/s)\eta_{s+r}} \left(M^{sg-(2s-r+1)h} [[K_{b,kb+h}(X)]]\right)^{r/s} + M^{-r/(3s)} X^{\eta_{s+r}}.$$

We are therefore led from the hypothesised bound (8.1) to the estimate

$$X^{\eta_{s+r}(1+\psi\theta)} \ll X^{(c+1)\delta} M^{-\gamma+rg-(2s-r+1)rh/s} (X/M^b)^{(1-r/s)\eta_{s+r}} [[K_{b,kb+h}(X)]]^{r/s} + X^{\eta_{s+r}-\delta},$$

whence

$$X^{\eta_{s+r}(r/s+(\psi+(1-r/s)b)\theta)} \ll X^{(c+1)\delta} M^{-\gamma+rg-(2s-r+1)rh/s} [[K_{b,kb+h}(X)]]^{r/s}.$$

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

Lemma 8.2. *Define s_0 and κ as in (2.6), and put $s = s_0$. Then $\eta_{s+r} = 0$.*

Proof. We begin by recalling our convention concerning the value of $K_{0,b}(X)$ from the preamble to Lemma 6.2. Thus, as a consequence of Lemma 6.1, it follows from (2.16) and (2.17) that there exists an integer h with $h \in \{0,1\}$ such that

$$[[J_{s+r}(X)]] \ll M^{-(2s-r+1)h} [[K_{0,1+h}(X)]].$$

We therefore deduce from (2.18) that, with $h = 0$ or 1, one has

$$X^{\eta_{s+r}} \ll X^\delta [[J_{s+r}(X)]] \ll X^\delta M^{-(2s-r+1)h} [[K_{0,1+h}(X)]]. \tag{8.2}$$

We may suppose that $\eta_{s+r} > 0$, for otherwise there is nothing to prove. We next take h_{-1} to be the integer h for which the relation (8.2) holds, and we define three sequences (a_n), (b_n), (h_n) of non-negative integers for $0 \leqslant n \leqslant N$ as follows. We put $a_0 = 0$ and $b_0 = 1 + h_{-1}$. Then, when $0 \leqslant n < N$, we fix any integer h_n with $0 \leqslant h_n \leqslant 2(k-1)b_n$, and then define

$$a_{n+1} = b_n \quad \text{and} \quad b_{n+1} = kb_n + h_n. \tag{8.3}$$

Next we define the auxiliary sequences (ψ_n), (c_n), (γ_n) of non-negative real numbers for $0 \leqslant n \leqslant N$ by putting $\psi_0 = 0$, $c_0 = 1$, $\gamma_0 = (2s-r+1)h_{-1}$. Then, for $0 \leqslant n < N$, we define

$$\psi_{n+1} = (s/r)\psi_n + (s/r-1)b_n, \tag{8.4}$$
$$c_{n+1} = (s/r)(c_n+1), \tag{8.5}$$
$$\gamma_{n+1} = (s/r)\gamma_n + (2s-r+1)h_n - sh_{n-1}. \tag{8.6}$$

We note that a straightforward induction reveals γ_n to be non-negative for $n \geqslant 0$, for the relation (8.6) yields the recurrence formula

$$\begin{aligned}\gamma_{n+1} - (2s-r+1)h_n &= (s/r)(\gamma_n - rh_{n-1})\\ &\geqslant (s/r)(\gamma_n - (2s-r+1)h_{n-1}).\end{aligned}$$

On recalling that $s/r = k$, we therefore see that for $n \geqslant 1$ one has

$$\begin{aligned}\gamma_n &\geqslant (2s-r+1)h_{n-1} + k^n(\gamma_0 - (2s-r+1)h_{-1})\\ &\geqslant (2s-r+1)h_{n-1} \geqslant 0,\end{aligned}$$

so that γ_n is indeed non-negative. A second induction confirms that for $0 \leqslant n \leqslant N$, one has

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

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

$$b_n < \sqrt{N}(s/r)^n \tag{8.7}$$

and

$$X^{\eta_{s+r}(1+\psi_n\theta)} \ll X^{c_n\delta}M^{-\gamma_n}[[K_{a_n,b_n}(X)]]. \tag{8.8}$$

When $n = 0$, the relation (8.7) holds by the definition of b_0. On the other hand, when $n = 0$, the relation (8.8) holds as a consequence of (8.2). We initiate further analysis of larger indices n with a preliminary discussion of the recurrence relations (8.3) to (8.6). Recall that $s = rk$, and observe that when $m \geqslant 1$, one has

$$\gamma_{m+1} - (s/r)\gamma_m = (2s-r+1)(b_{m+1} - kb_m) - s(b_m - kb_{m-1}),$$

whence

$$\gamma_{m+1} - (2s-r+1)b_{m+1} + sb_m = k(\gamma_m - (2s-r+1)b_m + sb_{m-1}).$$

It therefore follows by induction that for $m \geqslant 1$ one has

$$\gamma_m \geqslant (2s-r+1)b_m - sb_{m-1} + k^{m-1}(\gamma_1 - (2s-r+1)b_1 + sb_0).$$

We recall further that $b_0 = 1 + h_{-1}$, $b_1 = kb_0 + h_0$, and so

$$\begin{aligned}\gamma_1 - (2s-r+1)b_1 + sb_0 = (&k\gamma_0 + (2s-r+1)h_0 - sh_{-1})\\ &- (2s-r+1)(kb_0 + h_0) + sb_0.\end{aligned}$$

On recalling again the relation $s = rk$, we arrive at the formula

$$\begin{aligned}\gamma_1 - (2s-r+1)b_1 + sb_0 &= k(\gamma_0 - (2s-r+1)b_0) + s(b_0 - h_{-1})\\ &= s - rk - k(2s-2r+1),\end{aligned}$$

and this in turn delivers the lower bound

$$\gamma_m \geqslant (2s-r+1)b_m - sb_{m-1} - (2s-2r+1)k^m. \tag{8.9}$$

Suppose now that the desired conclusions (8.7) and (8.8) have been established for the index $n < N$. Then from (8.7), one has $kb_n\theta < k(s/r)^{n-N-2} < \frac{1}{3}$, whence $b_n < \frac{1}{3}(k\theta)^{-1}$. We may therefore appeal to Lemma 8.1 to deduce from

(8.8) that there exists a non-negative integer h, with $h \leqslant 2(k-1)b_n$, for which one has the upper bound

$$X^{\eta_{s+r}(1+\psi'\theta)} \ll X^{c'\delta} M^{-\gamma'} [[K_{a',b'}(X)]], \tag{8.10}$$

where

$$a' = b_n = a_{n+1}, \quad b' = kb_n + h, \tag{8.11}$$

$$\psi' = (s/r)\psi_n + (s/r-1)b_n = \psi_{n+1}, \tag{8.12}$$

$$c' = (s/r)(c_n+1) = c_{n+1}, \tag{8.13}$$

$$\gamma' = (s/r)\gamma_n + (2s-r+1)h - sh_{n-1}. \tag{8.14}$$

Notice here that in the final relation (8.14), we have made use of the formula $b_n - ka_n = b_n - kb_{n-1} = h_{n-1}$ available via (8.3).

Suppose, if possible, that $b' \geqslant \sqrt{N}(s/r)^{n+1} = \sqrt{N}k^{n+1}$. The relations (8.11) and (8.14) together with (8.9) show that

$$\begin{aligned}\gamma' &= (s/r)\gamma_n + (2s-r+1)(b'-kb_n) - s(b_n - kb_{n-1}) \\ &= k(\gamma_n - (2s-r+1)b_n + sb_{n-1}) + (2s-r+1)b' - sb_n \\ &\geqslant -(2s-2r+1)k^{n+1} + (2s-2r+1)b' + r(b'-kb_n) \\ &\geqslant (2s-2r+1)(b'-k^{n+1}) \geqslant (1-1/\sqrt{N})(2s-2r+1)b'. \end{aligned} \tag{8.15}$$

But $b' = kb_n + h \leqslant (3k-2)b_n < \theta^{-1}$, and so it follows from Lemma 5.3 that

$$[[K_{a',b'}(X)]] \ll X^{\eta_{s+r}+\delta}(M^{b'})^{\kappa}. \tag{8.16}$$

Combining (8.15), (8.16) and (8.10), therefore, we obtain the bound

$$X^{\eta_{s+r}(1+\psi_{n+1}\theta)} \ll X^{\eta_{s+r}+(c_{n+1}+1)\delta}(M^{b'})^{\kappa-(2s-2r+1)(1-1/\sqrt{N})}. \tag{8.17}$$

We now recall that $c_{n+1} \leqslant 3(s/r)^{n+1}$, so that $X^{(c_{n+1}+1)\delta} < M^{1/2}$. Also, when $r \geqslant 1$ and $k \geqslant 3$ one has

$$\begin{aligned}\kappa - &(1-1/\sqrt{N})(2s-2r+1) \\ &\leqslant (rk - \tfrac{1}{2}r(r+1))\left(\frac{k+1}{k-1}\right) - 2rk + 2r - 1 + 2s/\sqrt{N} \\ &\leqslant (rk-r)\left(\frac{k+1}{k-1}\right) + (2-2k)r - \tfrac{1}{2} \\ &= r(k+1) + (2-2k)r - \tfrac{1}{2} = (3-k)r - \tfrac{1}{2} \leqslant -\tfrac{1}{2}.\end{aligned}$$

Thus we obtain

$$X^{\eta_{s+r}(1+\psi_{n+1}\theta)} \ll X^{\eta_{s+r}} M^{(1-b')/2} \ll X^{\eta_{s+r}} M^{-1/2}. \tag{8.18}$$

Since ψ_{n+1} and θ are both positive, we are forced to conclude that $\eta_{s+r} < 0$, contradicting our opening hypothesis. The assumption that $b' \geqslant \sqrt{N}(s/r)^{n+1}$ is therefore untenable, and so we must in fact have $b' < \sqrt{N}(s/r)^{n+1}$. We take h_n to be the integer h at hand, so that $b' = b_{n+1}$ and $\gamma' = \gamma_{n+1}$, and thereby we obtain the desired conclusion that (8.7) and (8.8) hold with n replaced by $n+1$. This completes the present inductive step.

We have confirmed the validity of (8.7) and (8.8) for $0 \leqslant n \leqslant N$. We have also the bounds $c_n \leqslant 3(s/r)^n$, $\gamma_n \geqslant 0$ and $b_n \geqslant k^n$. Furthermore, since $s = rk$ one finds that

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

whence $\psi_n \geqslant n(k-1)k^{n-1}$. Finally, one has $b_N\theta < (r/s)^2 < 1$, so that $b_N < \theta^{-1}$. An application of Lemma 5.3 in combination with (8.8) therefore delivers the estimate

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

Again making use of the relation $\theta = N^{-1/2}(r/s)^{N+2}$ recorded in (2.9), we thus obtain the estimate

$$\eta_{s+r} \leqslant \frac{k^2}{\psi_N\theta} \leqslant \frac{\sqrt{N}k^2(s/r)^{N+2}}{N(k-1)k^{N-1}} < \frac{k^5}{\sqrt{N}}.$$

We are at liberty to take N as large as we please in terms of k, and thus η_{s+r} can be made arbitrarily small. It follows that $\eta_{s+r} = 0$, and this completes the proof of the lemma. □

The conclusion of Theorem 1.4 is an immediate consequence of Lemma 8.2. The latter shows that when $s \geqslant r(k+1)$, one has

$$J_s(X) \ll X^{2s-\kappa+\varepsilon},$$

where

$$\kappa = (rk - \tfrac{1}{2}r(r+1))\left(\frac{k+1}{k-1}\right).$$

Write $t = k - r$. Then this estimate may be rewritten to state that when $s \geqslant (k+1)(k-t)$, one has

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

where

$$\begin{aligned}\Delta_s &= \tfrac{1}{2}k(k+1) - \big((k-t)k - \tfrac{1}{2}(k-t)(k-t+1)\big)\left(\frac{k+1}{k-1}\right)\\ &= \tfrac{1}{2}t(t-1)\left(\frac{k+1}{k-1}\right).\end{aligned}$$

This completes the proof of Theorem 1.4 for $1 \leqslant t \leqslant k-1$. The special case in which $t = 1$ delivers the exponent $\Delta_s = 0$, so that when $s \geqslant k^2 - 1$ one has

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

The conclusion of Theorem 1.1 therefore follows as a speical case of Theorem 1.4.

9. The iterative process, II: quasi-diagonal behaviour

Our handling of the iterative process must be modified in order to establish Theorem 1.2, though the strategy is very similar to that underlying the proof of Theorem 1.4. There are sufficiently many differences from the treatment presented in §8 that, in the interests of enhancing clarity, we provide a fairly complete account in this section.

Lemma 9.1. *Suppose that* $1 \leqslant r \leqslant \min\{k-2, \frac{1}{2}k+1\}$*, define* s_0 *and* κ *as in (2.5), and put* $s = s_0$*. Let* a *and* b *be integers with* $b \geqslant (r-1)a$ *and* $0 \leqslant a < b \leqslant \frac{1}{3}(\rho\theta)^{-1}$ *having the property that when* $a = 0$*, one has* $b = 1$ *or* 2*, and put* $g = b - \rho a$*. Suppose in addition that there exist non-negative numbers* ψ*,* c *and* γ*, with* $c \leqslant 3(s/\rho)^N$*, for which*

$$X^{\eta_{s+r}(1+\psi\theta)} \ll X^{c\delta} M^{-\gamma} [[K_{a,b}(X)]]. \tag{9.1}$$

Then, for some non-negative integer h *with* $h \leqslant 2(\rho-1)b$*, one has*

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

where

$$\psi' = (s/r)\psi + (s/r-1)b, \quad c' = (s/r)(c+1),$$
$$a' = b, \quad b' = \rho b + h, \quad \gamma' = (s/r)\gamma + (2s-r+1)h - sg.$$

Proof. We follow the argument of the proof of Lemma 8.1, noting first that $X^{c\delta} < M^{1/(6s)}$ and $M^{1/(6s)} > X^{\delta}$. Then from Lemma 7.2 there exists an integer h with $0 \leqslant h \leqslant 2(\rho-1)b$ with the property that

$$\begin{aligned}[[K_{a,b}(X)]] \ll{}& X^{\delta}(X/M^b)^{(1-r/s)\eta_{s+r}} \left(M^{sg-(2s-r+1)h}[[K_{b,\rho b+h}(X)]]\right)^{r/s} \\ &+ M^{-r/(3s)} X^{\eta_{s+r}}.\end{aligned}$$

The hypothesised bound (9.1) therefore implies that

$$\begin{aligned}X^{\eta_{s+r}(1+\psi\theta)} \ll{}& X^{(c+1)\delta} M^{-\gamma+rg-(2s-r+1)rh/s} (X/M^b)^{(1-r/s)\eta_{s+r}} [[K_{b,\rho b+h}(X)]]^{r/s} \\ &+ X^{\eta_{s+r}-\delta},\end{aligned}$$

whence

$$X^{\eta_{s+r}(r/s+(\psi+(1-r/s)b)\theta)} \ll X^{(c+1)\delta} M^{-\gamma+rg-(2s-r+1)rh/s} [[K_{b,\rho b+h}(X)]]^{r/s}.$$

The conclusion of the lemma follows. □

Lemma 9.2. *Let* r *be a natural number with* $1 \leqslant r \leqslant \min\{k-2, \frac{1}{2}k+1\}$*. Define* s_0 *and* κ *as in (2.5), and put* $s = s_0$*. Then* $\eta_{s+r} = 0$*.*

Proof. We follow the proof of Lemma 8.2, supposing that $\eta_{s+r} > 0$. We begin by observing that the discussion of the first paragraph of the proof of Lemma 8.2 remains valid in the present circumstances, and so we may take h_{-1} to be an integer h for which the relation (8.2) holds. In this instance we define the sequences (a_n), (b_n), (h_n) of non-negative integers for $0 \leqslant n \leqslant N$ as follows. We put $a_0 = 0$ and $b_0 = 1 + h_{-1}$. Then, when $0 \leqslant n < N$, we fix any integer h_n with $0 \leqslant h_n \leqslant 2(\rho-1)b_n$, and then define

$$a_{n+1} = b_n \quad \text{and} \quad b_{n+1} = \rho b_n + h_n. \tag{9.2}$$

The auxiliary sequences (ψ_n), (c_n), (γ_n) of non-negative real numbers are defined for $0 \leqslant n \leqslant N$ by putting $\psi_0 = 0$, $c_0 = 1$, $\gamma_0 = (2s - r + 1)h_{-1}$. Then for $0 \leqslant n < N$, we define ψ_{n+1}, c_{n+1}, γ_{n+1} in terms of ψ_n, c_n, γ_n by means of the respective formulae (8.4), (8.5) and (8.6). We note that a straightforward induction again reveals γ_n to be non-negative for $n \geqslant 0$, just as in the proof of Lemma 8.2. One has $s/r = \rho$, and hence one finds that

$$\begin{aligned}\gamma_n &\geqslant (2s - r + 1)h_{n-1} + \rho^n(\gamma_0 - (2s - r + 1)h_{-1})\\ &\geqslant (2s - r + 1)h_{n-1} \geqslant 0.\end{aligned}$$

We also have $c_n \leqslant 3(s/r)^n$.

We claim that a choice may be made for the sequence (h_n) in such a manner that for $0 \leqslant n \leqslant N$, one has the upper bounds (8.7) and (8.8). As in our earlier discussion, these estimates hold for $n = 0$ as a consequence of the definition of b_0 together with (8.2). A comparison of the relations (9.2) and (8.3) reveals that the only adjustment necessary is to switch k in (8.3) to ρ in (9.2), though in present circumstances one has $s/r = \rho$. Thus we find as in the argument leading to (8.9) that in the present situation, one has for $m \geqslant 1$ that

$$\gamma_m \geqslant (2s - r + 1)b_m - sb_{m-1} - (2s - 2r + 1)\rho^m. \tag{9.3}$$

Suppose now that the desired conclusions (8.7) and (8.8) have been established for the index $n < N$. Then one has $\rho b_n \theta < \rho(s/r)^{n-N-2} < \frac{1}{3}$, whence $b_n < \frac{1}{3}(\rho\theta)^{-1}$. Also, our hypotheses on r ensure that

$$b_n \geqslant \rho b_{n-1} = (k - r + 1)a_n \geqslant (r - 1)a_n.$$

An application of Lemma 9.1 therefore leads from (8.8) to the conclusion that there exists an integer h, with $h \leqslant 2(\rho - 1)b_n$, for which one has the upper bound (8.10), where a', ψ', c', γ' satisfy (8.11)–(8.14), and in addition

$$b' = \rho b_n + h. \tag{9.4}$$

Suppose, if possible, that $b' \geqslant \sqrt{N}(s/r)^{n+1} = \sqrt{N}\rho^{n+1}$. Then as in the argument of the proof of Lemma 8.2 leading to (8.15) above, we find that (8.14) and (9.4) together with (9.3) show that

$$\begin{aligned}\gamma' &= (s/r)\gamma_n + (2s - r + 1)(b' - \rho b_n) - s(b_n - \rho b_{n-1})\\ &\geqslant (2s - 2r + 1)(b' - \rho^{n+1}) \geqslant (1 - 1/\sqrt{N})(2s - 2r + 1)b'.\end{aligned} \tag{9.5}$$

But $b' = \rho b_n + h \leqslant (3\rho - 2)b_n < \theta^{-1}$, and so it follows from Lemma 5.3 that (8.16) holds. Combining (9.5), (8.16) and (8.10), therefore, we obtain the bound (8.17). Observe next that in present circumstances, one deduces from (2.5) that

$$\begin{aligned}\kappa - (1 - 1/\sqrt{N})(2s - 2r + 1) &\leqslant s + r - \frac{r-1}{k-r} - 2s + 2r - 1 + \frac{2s}{\sqrt{N}}\\ &< 3r - \rho r - \tfrac{1}{2} = (r + 2 - k)r - \tfrac{1}{2}.\end{aligned}$$

Since, by assumption, we have $r \leqslant k - 2$, it follows that

$$\kappa - (1 - 1/\sqrt{N})(2s - 2r + 1) \leqslant -\tfrac{1}{2},$$

and thus we obtain again the relation (8.18). From here, one deduces as before that $\eta_{s+r} < 0$, contradicting our opening hypothesis, and leading us to conclude that in fact $b' < \sqrt{N}(s/r)^{n+1}$. We take h_n to be the integer h at hand, so that $b' = b_{n+1}$ and $\gamma' = \gamma_{n+1}$, and thereby deduce that (8.7) and (8.8) hold with n replaced by $n+1$. This completes the proof of the present inductive step.

Next, since (8.7) and (8.8) both hold for $0 \leqslant n \leqslant N$, one has $b_N\theta < (r/s)^2 < 1$, so that $b_N < \theta^{-1}$. From (9.2) one has $b_n \geqslant \rho^n$. Since $s = r\rho$, one finds that

$$\psi_{n+1} = \rho\psi_n + (\rho - 1)b_n \geqslant \rho\psi_n + (\rho - 1)\rho^n,$$

so that $\psi_n \geqslant n(\rho - 1)\rho^{n-1}$. An application of Lemma 5.3 therefore leads from (8.8) to the upper bound

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

But from (2.9) we have $\theta = N^{-1/2}(r/s)^{N+2}$, and thus

$$\eta_{s+r} \leqslant \frac{k^2}{\psi_N\theta} \leqslant \frac{\sqrt{N}k^2(s/r)^{N+2}}{N(\rho-1)\rho^{N-1}} < \frac{k^2\rho^3}{\sqrt{N}}.$$

On taking N sufficiently large in terms of k, we are able to make η_{s+r} as small as we please. It follows that $\eta_{s+r} = 0$, and this completes the proof of the lemma. □

The conclusion of Theorem 1.2 follows from Lemma 9.2. The latter shows that when $t = \rho r + r = r(k - r + 2)$, then one has

$$J_t(X) \ll X^{2t-(t-(r-1)/(k-r))+\varepsilon} = X^{t+\nu_t+\varepsilon},$$

in which $\nu_t = (r-1)/(k-r)$. When $s \leqslant t$, meanwhile, one may apply Hölder's inequality to obtain

$$\begin{aligned} J_s(X) = \oint |f_k(\boldsymbol{\alpha}; X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha} &\leqslant \Big(\oint |f_k(\boldsymbol{\alpha}; X)|^{2t}\,\mathrm{d}\boldsymbol{\alpha}\Big)^{s/t} \\ &\ll (X^{t+\nu_t+\varepsilon})^{s/t} \ll X^{s+\nu_t+\varepsilon}. \end{aligned}$$

This completes the proof of Theorem 1.2 for $1 \leqslant r \leqslant \min\{k-2, \frac{1}{2}k+1\}$.

We observe that when $k \geqslant 4$, the hypotheses of the statement of Theorem 1.2 are satisfied with $r = [(k+1)/2]$. In such circumstances, when $k = 2l+1$ is odd, one has

$$r(k-r+2) = (l+1)(l+2) \geqslant (l+\tfrac{1}{2})^2 + 2l + 1 = \tfrac{1}{4}k^2 + k,$$

and when $k = 2l$ is even, one has

$$r(k-r+2) = l(l+2) = \tfrac{1}{4}k^2 + k.$$

Meanwhile, one may easily verify that in each case the exponent $\nu_{r,k}$ satisfies

$$\nu_{r,k} = \frac{r-1}{k-r} \leqslant 1.$$

The conclusion of Corollary 1.3 therefore follows directly from Theorem 1.2.

Finally, suppose that $2 \leqslant r \leqslant \min\{k-2, \frac{1}{2}k+1\}$, and put $t(r) = r(k-r+2)$. Then whenever $t(r-1) \leqslant s \leqslant t(r)$, it is a consequence of Theorem 1.2 that $J_{s,k}(X) \ll X^{s+\nu+\varepsilon}$, where

$$\nu = \frac{r-1}{k-r} \leqslant \frac{t(r-1)}{(k-r)(k-r+3)} \leqslant \frac{4s}{k^2}.$$

Thus we see that the upper bound (1.3) does indeed hold with a permissible exponent $\delta_{s,k}$ satisfying $\delta_{s,k} = O(s/k^2)$, thereby justifying the discussion following the statement of Theorem 1.2.

10. The asymptotic formula in Waring's problem

Our first application of the improved mean value estimate supplied by Theorem 1.1 concerns the asymptotic formula in Waring's problem. In this context, we define the exponential sum $g(\alpha) = g_k(\alpha; X)$ by

$$g_k(\alpha; X) = \sum_{1 \leqslant x \leqslant X} e(\alpha x^k).$$

Also, we define the set of minor arcs $\mathfrak{m} = \mathfrak{m}_k$ to be the set of real numbers $\alpha \in [0, 1)$ satisfying the property that, whenever $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $(a, q) = 1$ and $|q\alpha - a| \leqslant (2k)^{-1}X^{1-k}$, then $q > (2k)^{-1}X$. We begin by applying the methods of [21] to derive a mean value estimate restricted to minor arcs.

Theorem 10.1. *Suppose that $s \geqslant k^2 - 1$. Then for each $\varepsilon > 0$, one has*

$$\int_{\mathfrak{m}} |g_k(\alpha; X)|^{2s} \, \mathrm{d}\alpha \ll X^{2s-k-1+\varepsilon}.$$

Proof. According to [21, Theorem 2.1], one has

$$\int_{\mathfrak{m}} |g_k(\alpha; X)|^{2s} \, \mathrm{d}\alpha \ll X^{\frac{1}{2}k(k-1)-1} (\log X)^{2s+1} J_{s,k}(2X).$$

Theorem 1.1 shows that when $s \geqslant k^2 - 1$, one has $J_{s,k}(2X) \ll X^{2s-\frac{1}{2}k(k+1)+\varepsilon}$, and the conclusion of the theorem now follows. □

We transform the estimate supplied by this theorem into a less strident bound useful in handling the minor arc contribution in Waring's problem. For each natural number k, define the positive integer $s_0(j) = s_0(k, j)$ by means of the relation

$$s_0(k, j) = 2k^2 - 2k - \frac{2(k-1)(j+1) - 2^{j+1}}{k-j}.$$

We then put

$$s_1(k) = \min_{\substack{0 \leqslant j \leqslant k-2 \\ 2^j \leqslant k^2-k-1}} s_0(k, j). \tag{10.1}$$

Lemma 10.2. *Suppose that k is a natural number with $k \geqslant 3$. Then*

$$\int_0^1 |g_k(\alpha;X)|^{s_1(k)}\,\mathrm{d}\alpha \ll X^{s_1(k)-k+\varepsilon}.$$

Moreover, when s is a real number with $s > s_1(k)$, there exists a positive number $\delta = \delta(k,s)$ with the property that

$$\int_{\mathfrak{m}} |g_k(\alpha;X)|^s\,\mathrm{d}\alpha \ll X^{s-k-\delta}.$$

Proof. The second estimate claimed in the lemma is immediate from Theorem 10.1 when $s \geqslant 2k^2-2$, on making use of the trivial estimate $|g_k(\alpha;X)| \leqslant X$. We suppose therefore that $s_1(k) < s \leqslant 2k^2-2$, and we put $\tau = s - s_1(k)$. Let j be an integer with $0 \leqslant j \leqslant k-2$ and $2^j \leqslant k(k-1)-1$ for which $s_1(k) = s_0(k,j)$. Then by Hölder's inequality, one has

$$\int_{\mathfrak{m}} |g(\alpha)|^s\,\mathrm{d}\alpha \leqslant \Big(\int_{\mathfrak{m}} |g(\alpha)|^{2k^2-2}\,\mathrm{d}\alpha\Big)^a \Big(\int_0^1 |g(\alpha)|^{2^{j+1}}\,\mathrm{d}\alpha\Big)^b,$$

where

$$a = \frac{s-2^{j+1}}{2k^2-2-2^{j+1}} \quad \text{and} \quad b = \frac{2k^2-2-s}{2k^2-2-2^{j+1}}.$$

An application of Theorem 10.1 in combination with Hua's lemma (see [13, Lemma 2.5]) therefore yields the bound

$$\begin{aligned}\int_{\mathfrak{m}} |g(\alpha)|^s\,\mathrm{d}\alpha &\ll X^\varepsilon (X^{(2k^2-2)-k-1})^a (X^{2^{j+1}-j-1})^b \\ &\ll X^{s-k-\nu+\varepsilon},\end{aligned}$$

where $\nu = a-(k-j-1)b$. A modicum of computation reveals that

$$\nu = \frac{(k-j)(s-s_1(k))}{2k^2-2-2^{j+1}} \geqslant \tau/(2k^2),$$

and so the second conclusion of the lemma therefore follows with $\delta = \tau/(4k^2)$.

When $s = s_1(k)$, the above discussion shows that

$$\int_{\mathfrak{m}} |g(\alpha)|^s\,\mathrm{d}\alpha \ll X^{s-k+\varepsilon}. \tag{10.2}$$

But on writing $\mathfrak{M} = [0,1) \setminus \mathfrak{m}$, the methods of [13, Chapter 4] confirm that whenever $s \geqslant k+2$, one has

$$\int_{\mathfrak{M}} |g(\alpha)|^s\,\mathrm{d}\alpha \ll X^{s-k}.$$

The first conclusion of the lemma follows by combining this estimate with the earlier bound (10.2). □

The argument following the proof of [21, Lemma 3.1] may now be adapted, without effort, to show that $\widetilde{G}(k) \leqslant [s_1(k)]+1$ for $k \geqslant 3$. The first conclusion of Theorem 1.5 consequently follows at once from the definition (10.1). This upper bound for $\widetilde{G}(k)$ is easily made explicit for smaller values of k. Thus, on taking $r=3$, one finds that for $k \geqslant 5$ one has

$$\frac{2(k-1)(r+1)-2^{r+1}}{k-r} = \frac{8k-24}{k-3} = 8,$$

and on taking $r=4$, one finds that for $k \geqslant 6$ one has

$$\frac{2(k-1)(r+1)-2^{r+1}}{k-r} = \frac{10k-42}{k-4} = 10 - \frac{2}{k-4},$$

which is at least 9 for $k \geqslant 6$, and exceeds 9 for $k \geqslant 7$. Also, on taking $r=5$, one finds that

$$\frac{2(k-1)(r+1)-2^{r+1}}{k-r} = \frac{12k-76}{k-5} = 12 - \frac{16}{k-5},$$

a quantity which exceeds 10 for $k \geqslant 14$. Thus we deduce that

$$\widetilde{G}(6) \leqslant 52, \quad \widetilde{G}(k) \leqslant 2k^2-2k-9 \quad (7 \leqslant k \leqslant 13)$$

and

$$\widetilde{G}(k) \leqslant 2k^2-2k-10 \quad (k \geqslant 14).$$

An alternative to the above approach proceeds by means of the methods of Ford [6]. Motivated by the notation introduced in (2.16), we write

$$[[J_{t,k}(Y)]]^* = Y^{\frac{1}{2}k(k+1)-2t} J_{t,k}(Y).$$

One may then rephrase [6, Theorem 1] in the following form.

Theorem 10.3. *Let m be an integer with $1 \leqslant m \leqslant k$. Then for each natural number s with $s \geqslant \frac{1}{2}m(m-1)$, one has*

$$\int_0^1 |g_k(\alpha;X)|^{2s}\,\mathrm{d}\alpha \ll X^{2s-k}[[J_{s-\frac{1}{2}m(m-1),k}(X^{1/m})]]^*.$$

For each natural number k, we now consider integers m and t with $1 \leqslant m \leqslant k$ and $1 \leqslant t \leqslant k-1$, and we define $\Delta_{t,k}$ as in (1.4). We then put

$$s_2(k,m,t) = 2k^2 - 2 - \frac{2(t-1)(k+1)-m(m-1)}{1+\Delta_{t,k}/m},$$

and set

$$s_3(k) = \min_{\substack{1\leqslant m\leqslant k \\ 2(t-1)(k+1)+m(m-1)<2k^2-2}} \min_{1\leqslant t\leqslant k-1} s_2(k,m,t).$$

Lemma 10.4. *Suppose that s and k are natural numbers with $k \geqslant 3$ and $s > s_3(k)$. Then there exists a positive number $\delta = \delta(k,s)$ with the property that*

$$\int_{\mathfrak{m}} |g_k(\alpha;X)|^s\,\mathrm{d}\alpha \ll X^{s-k-\delta}.$$

Proof. As in the proof of Lemma 10.2, the desired conclusion is immediate from Theorem 10.1 when $s \geqslant 2k^2-2$, on making use of the trivial estimate $|g_k(\alpha;X)| \leqslant X$. We suppose therefore that $s_3(k) < s \leqslant 2k^2-2$, and we put $\tau = s - s_3(k)$. Let m and t be integers with $1 \leqslant m \leqslant k$, $1 \leqslant t \leqslant k-1$ and $2(t-1)(k+1)+m(m-1) < 2k^2-2$, for which $s_3(k) = s_2(k,m,t)$. Then by Hölder's inequality, one has

$$\int_{\mathfrak{m}} |g(\alpha)|^s \, \mathrm{d}\alpha \leqslant \Bigl(\int_{\mathfrak{m}} |g(\alpha)|^{2k^2-2} \, \mathrm{d}\alpha \Bigr)^a \Bigl(\int_0^1 |g(\alpha)|^{2(k-t)(k+1)+m(m-1)} \, \mathrm{d}\alpha \Bigr)^b,$$

where

$$a = \frac{s-2(k-t)(k+1)-m(m-1)}{2k^2-2-2(k-t)(k+1)-m(m-1)}$$

and

$$b = \frac{2k^2-2-s}{2k^2-2-2(k-t)(k+1)-m(m-1)}.$$

By applying Theorem 10.3 and Theorem 1.4 in sequence, one finds that

$$\int_0^1 |g(\alpha)|^{2(k-t)(k+1)+m(m-1)} \, \mathrm{d}\alpha \ll X^{2(k-t)(k+1)+m(m-1)-k+\Delta_{t,k}/m+\varepsilon}.$$

Consequently, an application of Theorem 10.1 yields the bound

$$\begin{aligned} \int_{\mathfrak{m}} |g(\alpha)|^s \, \mathrm{d}\alpha &\ll X^\varepsilon (X^{(2k^2-2)-k-1})^a (X^{2(k-t)(k+1)+m(m-1)-k+\Delta_{t,k}/m})^b \\ &\ll X^{s-k-\nu+\varepsilon}, \end{aligned}$$

where

$$\nu = a - b\Delta_{t,k}/m = \frac{(1+\Delta_{t,k}/m)\,(s-s_3(k))}{2k^2-2-2(k-t)(k+1)-m(m-1)} \geqslant \tau/(2k^2).$$

The conclusion of the lemma therefore follows with $\delta = \tau/(4k^2)$. □

The argument following the proof of [21, Lemma 3.1] may again be adapted to show that $\widetilde{G}(k) \leqslant [s_3(k)]+1$ for $k \geqslant 3$. One can check by means of a direct computation that when $k=20$, if one takes $t=7$ and $m=9$, then $s_2(k,m,t) < 748$, and in this way one obtains the bound $\widetilde{G}(20) \leqslant 748$. In view of the discussion following the proof of Lemma 10.2, this completes the proof of Corollary 1.7. Similarly, the conclusion of Corollary 1.6 follows on taking $t = 2[k^{1/3}]$ and $m = [k^{2/3}]$, for then one finds that

$$\Delta_{t,k}/m = \frac{\frac{1}{2}t(t-1)}{m} \left(\frac{k+1}{k-1} \right) = \frac{2k^{2/3}+O(k^{1/3})}{k^{2/3}+O(1)} = 2 + O(k^{-1/3}),$$

and hence

$$\begin{aligned} s_2(k,m,t) &= 2k^2 - 2 - \frac{2k(2k^{1/3}) - k^{4/3} + O(k)}{3+O(k^{-1/3})} \\ &= 2k^2 - k^{4/3} + O(k). \end{aligned}$$

We finish by noting that the proof of [21, Theorem 4.2] may be adapted transparently so as to establish that when $s > \min\{s_1(k), s_3(k)\}$, then the anticipated asymptotic formula holds for the number of integral solutions of the diagonal equation

$$a_1x_1^k+\ldots+a_sx_s^k=0,$$

with $|\mathbf{x}|\leqslant B$. Here, the coefficients a_i $(1\leqslant i\leqslant s)$ are fixed integers. Similar improvements may be wrought in upper bounds for $\widetilde{G}^+(k)$, the least number of variables required to establish that the anticipated asymptotic formula in Waring's problem holds for almost all natural numbers n. Thus, one may adapt the methods of [21, §5] to show that

$$\widetilde{G}^+(k)\leqslant k^2-k+1-\max_{\substack{0\leqslant j\leqslant k-2\\ 2^j\leqslant k^2-k-1}}\left\lceil \frac{(k-1)(j+1)-2^j}{k-j}\right\rceil$$

and

$$\widetilde{G}^+(k)\leqslant k^2-\max_{\substack{1\leqslant m\leqslant k\\ 2(t-1)(k+1)+m(m-1)<2k^2-2}}\max_{1\leqslant t\leqslant k-1}\left\lceil \frac{(t-1)(k+1)-\frac{1}{2}m(m-1)}{1+\Delta_{t,k}/m}\right\rceil .$$

11. Further applications

In this section we briefly discuss some applications of the mean value estimates supplied by Theorems 1.1 and 1.4, with the aim of noting improvements made available over our previous work [20]. We begin with an analogue of Weyl's inequality.

Theorem 11.1. *Let k be an integer with $k\geqslant 4$, and let $\boldsymbol{\alpha}\in\mathbb{R}^k$. Suppose that there exists a natural number j with $2\leqslant j\leqslant k$ such that, for some $a\in\mathbb{Z}$ and $q\in\mathbb{N}$ with $(a,q)=1$, one has $|\alpha_j-a/q|\leqslant q^{-2}$ and $q\leqslant X^j$. Then one has*

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

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

Proof. Under the hypotheses of the statement of the theorem, we find that [13, Theorem 5.2] shows that for $s\in\mathbb{N}$, one has

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

The conclusion of the theorem therefore follows on taking

$$s=(k-1)^2-1=k(k-2),$$

for in such circumstances Theorem 1.1 delivers the bound

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

□

The proof of [20, Theorem 1.6] may be easily adapted to deliver estimates depending on common diophantine approximations.

Theorem 11.2. *Let k be an integer with $k\geqslant 4$, and let τ and δ be real numbers with $\tau^{-1}>4k(k-2)$ and $\delta>k\tau$. Suppose that X is sufficiently large in terms of k, δ and τ, and further that $|f_k(\boldsymbol{\alpha};X)|>X^{1-\tau}$. Then there exist integers $q,a_1,\ldots,a_k$ such that $1\leqslant q\leqslant X^\delta$ and $|q\alpha_j-a_j|\leqslant X^{\delta-j}$ $(1\leqslant j\leqslant k)$.*

The proof of [20, Theorem 1.7] likewise delivers the following result concerning the distribution modulo 1 of polynomial sequences. Here, we write $\|\theta\|$ for $\min_{y\in\mathbb{Z}}|\theta - y|$.

Theorem 11.3. *Let k be an integer with $k \geqslant 4$, and define $\tau(k)$ by $\tau(k)^{-1} = 4k(k-2)$. Then whenever $\boldsymbol{\alpha} \in \mathbb{R}^k$ and N is sufficiently large in terms of k and ε, one has*

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

In each of Theorems 11.2 and 11.3, the exponent $4k(k-2)$ represents an improvement on the exponent $4k(k-1)$ made available in [20, Theorems 1.6 and 1.7]. In [20, Theorem 1.5], meanwhile, we established a conclusion similar to that of Theorem 11.1, though with a weaker exponent $\sigma(k)$ satisfying $\sigma(k)^{-1} = 2k(k-1)$. As with this earlier work, our estimates supersede the Weyl exponent $\sigma(k) = 2^{1-k}$ when $k \geqslant 8$, and supersede work of Heath-Brown [7] and Robert and Sargos [10] for $k \geqslant 9$. When $k = 8$, in fact, our exponent matches that of Heath-Brown [7], though our conclusion is applicable for a substantially larger set of coefficients.

We turn next to Tarry's problem. When h, k and s are positive integers with $h \geqslant 2$, consider the Diophantine system

$$\sum_{i=1}^{s} x_{i1}^j = \sum_{i=1}^{s} x_{i2}^j = \ldots = \sum_{i=1}^{s} x_{ih}^j \quad (1 \leqslant j \leqslant k). \tag{11.1}$$

Let $W(k,h)$ denote the least natural number s having the property that the simultaneous equations (11.1) possess an integral solution $\mathbf{x}$ with

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

Theorem 11.4. *When h and k are natural numbers with $h \geqslant 2$ and $k \geqslant 2$, one has $W(k,h) \leqslant k^2 - \sqrt{2}k^{3/2} + 4k$.*

Proof. The argument of the proof of [20, Theorem 1.3] shows that $W(k,h) \leqslant s$ whenever $J_{s,k+1}(X) = o(J_{s,k}(X))$. Incorporating the bounds for $J_{s,k+1}(X)$ supplied via Theorem 1.4 into this argument, one finds that

$$W(k,h) \leqslant (k+1-t)(k+2)$$

whenever

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

a constraint equivalent to the condition

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

By direct computation, one finds that this inequality is satisfied when $t = [\sqrt{2k}]$, but not for $t \geqslant \sqrt{2k}+1$. Thus we deduce that

$$\begin{aligned} W(k,h) &\leqslant (k-[\sqrt{2k}]+1)(k+2) = k^2+3k-(k+2)[\sqrt{2k}]+2 \\ &\leqslant k^2-\sqrt{2}k^{3/2}+4k+4-2\sqrt{2k}. \end{aligned}$$

The conclusion of the theorem follows immediately. □

In [20, Theorem 1.3], we obtained the weaker bound $W(k,h) \leqslant k^2+k-2$. We remark that the conclusion of Theorem 11.4 may be utilised to obtain an improvement in a result of Croot and Hart related to the sum-product theorem. When A is a set of real numbers, write

$$A \cdot A = \{xy : x \in A \text{ and } y \in A\}$$

and

$$hA = \{x_1+\ldots+x_h : x_i \in A\ (1 \leqslant i \leqslant h)\}.$$

Theorem 11.5. *Suppose that h and n are natural numbers with $h \geqslant 2$. Let A be a set of n real numbers. Then whenever ε is a positive number sufficiently small in terms of h, and $|A\cdot A| \leqslant n^{1+\varepsilon}$, there exists a positive number λ having the property that*

$$|h(A\cdot A)| > n^{\lambda h^{1/3}}.$$

The aforementioned result of Croot and Hart (see [5, Theorem 1.2]) delivers a similar conclusion, though with the exponent $h^{1/3}$ replaced by $(h/\log h)^{1/3}$.

We note also that on writing

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

and

$$\mathcal{J}(s,k) = \int_{\mathbb{R}^k} \Bigl| \int_0^1 e(\beta_1\gamma+\ldots+\beta_k\gamma^k)\,\mathrm{d}\gamma \Bigr|^{2s} \mathrm{d}\boldsymbol{\beta},$$

the method of proof of [20, Theorem 1.2] may be modified in the light of Theorem 1.1 to obtain the asymptotic formula

$$J_{s,k}(X) \sim \mathfrak{S}(s,k)\mathcal{J}(s,k)X^{2s-\frac{1}{2}k(k+1)},$$

provided only that $k \geqslant 3$ and $s \geqslant k^2$. In [20, Theorem 1.2], such a conclusion was obtained for $s \geqslant k^2+k+1$. A similar improvement holds also for work on the asymptotic formula in the Hilbert-Kamke problem.

Finally, write

$$F_k(\boldsymbol{\beta};X) = \sum_{1\leqslant x\leqslant X} e(\beta_kx^k+\beta_{k-2}x^{k-2}+\ldots+\beta_1x).$$

L.-K. Hua investigated the problem of bounding the least integer C_k such that, whenever $s \geqslant C_k$, one has

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

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

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

Theorem 11.6. *When $k \geqslant 3$, one has $C_k \leqslant 2k^2 - 2$ and $S_k \leqslant 2k^2 - 2k$.*

Proof. The bound on C_k is immediate from Theorem 1.1. In order to establish the bound on S_k, we begin by observing that [20, equation (10.10)] supplies the estimate

$$\oint |F_k(\boldsymbol{\beta}; X)|^{2t} \, \mathrm{d}\boldsymbol{\beta} \ll X^{k-2+\varepsilon} J_{t,k}(2X) + X^{\varepsilon-1} J_{t,k-1}(2X). \tag{11.2}$$

Write $u = (k-2)(k+1)$. Then an application of Theorem 1.4 with $t = 2$ shows that

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

with $\Delta = (k+1)/(k-1)$. Consequently, on applying Hölder's inequality in combination with Theorem 1.1, we obtain the bound

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

On the other hand, it follows from Theorem 1.1 that whenever $s \geqslant k(k-2)$, then one has

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

On substituting these estimates into (11.2), we conclude that

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

We therefore see that $S_k \leqslant 2k(k-1)$, and this completes the proof of the theorem. □

For comparison, in [20, Theorems 1.1 and 10.3] we derived the weaker bounds $C_k \leqslant 2k^2 + 2k$ and $S_k \leqslant 2k^2 + 2k - 4$. When $k \geqslant 4$, the conclusion of Theorem 11.6 improves also on the bounds obtained by Hua [8, Chapter 5], namely

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

and

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

Moreover, Theorem 11.6 matches the bound established by Hua for C_3.

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] G. I. Arkhipov and A. A. Karatsuba, *A new estimate of an integral of I. M. Vinogradov*, Izv. Akad. Nauk SSSR Ser. Mat. **42** (1978), 751–762.

[3] K. D. Boklan, *The asymptotic formula in Waring's problem*, Mathematika **41** (1994), 329–347.

[4] K. D. Boklan and T. D. Wooley, *On Weyl sums for smaller exponents*, Funct. Approx. Comment. Math. (to appear).

[5] E. Croot and D. Hart, *h-fold sums from a set with few products*, SIAM J. Discrete Math. **24** (2010), 505–519.

[6] K. B. Ford, *New estimates for mean values of Weyl sums*, Internat. Math. Res. Notices (1995), 155–171.

[7] D. R. Heath-Brown, *Weyl's inequality, Hua's inequality, and Waring's problem*, J. London Math. Soc. (2) **38** (1988), 216–230.

[8] L.-K. Hua, *Additive theory of prime numbers*, American Math. Soc., Providence, RI, 1965.

[9] S. T. Parsell, *On the Bombieri-Korobov estimate for Weyl sums*, Acta Arith. **138** (2009), 363–372.

[10] O. Robert and P. Sargos, *Un théorème de moyenne pour les sommes d'exponentielles. Application á l'inégalité de Weyl*, Publ. Inst. Math. (Beograd) (N.S.) **67** (2000), 14–30.

[11] O. V. Tyrina, *A new estimate for a trigonometric integral of I. M. Vinogradov*, Izv. Akad. Nauk SSSR Ser. Mat. **51** (1987), 363–378.

[12] R. C. Vaughan, *On Waring's problem for smaller exponents, II*, Mathematika **33** (1986), 6–22.

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

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

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

[16] T. D. Wooley, *On Vinogradov's mean value theorem*, Mathematika **39** (1992), 379–399.

[17] T. D. Wooley, *Quasi-diagonal behaviour in certain mean value theorems of additive number theory*, J. Amer. Math. Soc. **7** (1994), 221–245.

[18] T. D. Wooley, *Some remarks on Vinogradov's mean value theorem and Tarry's problem*, Monatsh. Math. **122** (1996), 265–273.

[19] T. D. Wooley, *A note on simultaneous congruences*, J. Number Theory **58** (1996), 288–297.

[20] T. D. Wooley, *Vinogradov's mean value theorem via efficient congruencing*, Annals of Math. (to appear), arXiv:1101.0574.

[21] T. D. Wooley, *The asymptotic formula in Waring's problem*, Internat. Math. Res. Notices (in press).

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

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