Rational points on linear slices of diagonal hypersurfaces

Jörg Brüdern, Olivier Robert

RATIONAL POINTS ON LINEAR SLICES OF DIAGONAL HYPERSURFACES

JÖRG BRÜDERN AND OLIVIER ROBERT

ABSTRACT. An asymptotic formula is obtained for the number of rational points of bounded height on the class of varieties described in the title line. The formula is proved via the Hardy-Littlewood method, and along the way, we establish two new results on Weyl sums that are of some independent interest.

1. INTRODUCTION

The varieties alluded to in the title line are defined by pairs of equations

$$(1.1) \qquad \sum_{j=1}^{s} a_j x_j^k = \sum_{j=1}^{s} b_j x_j = 0$$

in which the natural numbers k, s and the integers a_j, b_j are fixed once and for all. We shall be concerned with deriving an asymptotic formula for the number $N(P) = N_{\mathbf{a},\mathbf{b}}(P)$ of solutions to (1.1) in integers x_j satisfying

$$(1.2) \qquad |x_j| \le P \quad (1 \le j \le s).$$

The cases $k = 1$ and $k = 2$ are part of the classical theory: When $k = 1$, the equations (1.1) describe a lattice, and the asymptotic evaluation of $N(P)$ is elementary. When $k = 2$, one inserts the linear equation into the quadratic one to eliminate a variable, thus reducing the problem to that of counting those integer points where an integral quadratic form vanishes. For the latter problem, there is a vast literature to which we have nothing to add. Thus, we concentrate on the cases where $k \ge 3$.

Theorem 1. *Let $k \ge 3$, $s \ge 2^k + 2$ and suppose that $a_j \ne 0$ ($1 \le j \le s$). Suppose that the pair of equations (1.1) has non-singular solutions in $\mathbb{R}$ and in $\mathbb{Q}_p$, for all primes p. Then there is a positive number $\mathfrak{C}(\mathbf{a}, \mathbf{b})$ such that*

$$(1.3) \qquad N_{\mathbf{a},\mathbf{b}}(P) = \mathfrak{C}(\mathbf{a}, \mathbf{b}) P^{s-k-1} + O\big(P^{s-k-1}(\log P)^{-2}\big).$$

In algebraic geometry, it is more customary to count rational points on the projective variety defined by (1.1). A rational point on (1.1) corresponds to an integral solution with $(x_1; x_2; \ldots; x_s) = 1$. The latter is unique up to sign, and its natural height is defined by $\max |x_j|$. By Möbius' inversion formula, the number of rational points on (1.1) with height not exceeding P equals

$$\frac{1}{2} \sum_{d \le P} \mu(d)\big(N(P/d) - 1\big).$$

Subject to the conditions in Theorem 1, this expression is asymptotic to

$$\tfrac{1}{2}\zeta(s-k-1)^{-1}\mathfrak{C}(\mathbf{a},\mathbf{b})P^{s-k-1},$$

as expected.

Theorem 1 should be considered as part of a programme to establish similar asymptotic formulae for intersections of diagonal hypersurfaces, at least when the dimension is suitably large (see [6, 9, 10, 23, 32, 33] and the references therein). Recent ground-breaking work of Wooley [35, 37, 36] on Vinogradov's mean value theorem has a revolutionary impact in this area. At the time of writing, publicly available descriptions of Wooley's "efficient congruencing" provide conclusions similar to Theorem 1, but subject to a condition slightly milder than $s \geq 2(k-1)^2$ when $k \geq 6$. However, progress is still ongoing, and Wooley has now announced results that have the potential to supersede Theorem 1 for all $k \geq 5$. In the light of this, the main interest is in the cases $k = 3$ and $k = 4$, but our proof for $k = 4$ works equally well for $k \geq 4$.

When $k = 3$, the condition on s in Theorem 1 is $s \geq 10$. As in the quadratic case, one may substitute the linear equation into the cubic one to obtain a cubic form in nine or more variables. For cubic forms in nine variables, important work of Hooley [18, 19, 20] provides an asymptotic formula for the number of its integral zeros within a suitable expanding region, provided that the form is non-singular, a condition that may in some cases be relaxed to allow the singular locus of the form to consist of isolated linearly independent ordinary double points. However, as one readily checks, the projective cubic defined by

$$\sum_{j=1}^{10} x_j^3 = \sum_{j=1}^{10} x_j = 0$$

has 126 singular points, and hence provides an example covered by Theorem 1 but not by Hooley's work. When $k = 4$, one may again insert the linear equation into the quartic one. This leads to a quartic form that one may analyse by the methods of Birch [1] and Browning and Heath-Brown [4], but this strategy apparently requires s to be as large as 40 or thereabouts, and is therefore is not competitive at present.

We prove Theorem 1 by a two-dimensional version of the Hardy-Littlewood method. Our argument rests on a new mean value theorem for the generating function

$$(1.4) \qquad f(\alpha,\beta) = \sum_{x \leq P} e(\alpha x^k + \beta x)$$

that we now describe. Fix a number θ with $2^{1-k}k < \theta \leq \tfrac{4}{5}$ and then take $Q = P^\theta$. Let $\mathfrak{m}$ denote the set of real numbers $\alpha \in [0,1]$ for which the inequality $|q\alpha - a| \leq QP^{-k}$ with $q \in \mathbb{N}$, $a \in \mathbb{Z}$ is only possible when $q > Q$.

Theorem 2. *Let $k \geq 3$. Then*

$$\int_0^1 \int_{\mathfrak{m}} |f(\alpha,\beta)|^{2^k+2}\,\mathrm{d}\alpha\,\mathrm{d}\beta \ll P^{2^k-k+1}(\log P)^{-2}.$$

Theorem 2 should be compared with the celebrated estimate

$$(1.5) \qquad \int_{\mathfrak{m}} |f(\alpha,0)|^{2^k}\,\mathrm{d}\alpha \ll P^{2^k-k}(\log P)^{-2}$$

due to Vaughan (see [26], Theorem B, for $k = 3$ and [27] for $k \geq 4$, with a slightly different power of $\log P$; for refinements see Boklan [2] and Harvey [14]). Vaughan's approach involves a considerable refinement of the conventional proof of Hua's Lemma (see [29], Lemma 2.5). Certain divisor sums require treatment by mean value estimates for Hooley's Δ_r-functions to propagate an initial saving through an induction. The initial saving itself comes in through a sieving of the variable of summation in (1.4) and an appeal to a paucity estimate for the diophantine equation

$$x_1^k + x_2^k = x_3^k + x_4^k.$$

originally obtained by Hooley [15, 17, 21] (see also Wooley [34], Skinner and Wooley [24]). We are able to keep the architecture of Vaughan's treatment largely intact, now building on a paucity estimate for the pair

$$x_1^k + x_2^k + x_3^k = y_1^k + y_2^k + y_3^k, \quad x_1 + x_2 + x_3 = y_1 + y_2 + y_3$$

(Vaughan and Wooley [31] for $k = 3$, Greaves [11] for $k \geq 4$; see also de la Bretèche [3]). Once Theorem 2 is established, it is fairly routine to derive Theorem 1. We postpone a more comprehensive discussion of several complications to the appropriate stage of the argument.

No direct attack on the problem considered here, via the circle method, seems to have been launched in the past, but there is related work of Parsell [23]. Parsell considers more generally a pair of diagonal equations

$$a_1 x_1^k + \ldots + a_s x_s^k = b_1 x_1^n + \ldots + b_s x_s^n = 0,$$

and applies smooth number technology within a circle method approach to verify the Hasse principle for this pair of equations when s is suitably large. Such a strategy typically supplies a lower bound for the number of solutions within a box which is of the expected order of magnitude. In the special case $n = 1$ which is the theme of this paper, Parsell proves the Hasse principle for $k = 3$, $s \geq 10$ and for $k = 4$, $s \geq 17$, amongst other results. It is interesting to note that in the case $k = 3$ his method fails to give a lower bound for $N(P)$ of the expected size, a defect that is now cured by Theorem 1.

Before we move on to proofs of Theorem 1 and 2, we briefly comment on the condition in Theorem 1 that all a_j be non-zero. It suffices to require only that at least 2^k of the a_j are non-zero. In fact, the presence of isolated linear variables in (1.1) facilitates the exercise. However, it seems difficult to relax this condition further without improving (1.5), and some lower bound on the number of non-zero a_j is definitely necessary. To see this, consider the system

(1.6) $$5x_1^3 + 9x_2^3 + 10x_3^3 + 12x_4^3 = x_1 + x_2 + \cdots + x_{10} = 0.$$

Here, the cubic equation (when considered in the variables $x_1, \ldots, x_4$ only) violates the Hasse principle (Cassels and Guy [7]). Thus there are real and p-adic non-singular solutions of (1.6), but a solution $\mathbf{x} \in \mathbb{Z}^{10}$ satisfies $x_1 = x_2 = x_3 = x_4 = 0$, and it follows easily that there is a positive constant C such that $N(P) = CP^5 + O(P^4)$, in contrast to the leading term of size P^{s-k-1} in Theorem 1.

Notation. Throughout this paper, small italics $a, b, \ldots$ denote integers, and q is a natural number. The letter p is reserved for primes, and k is a natural number with $k \geq 3$. Real numbers are denoted by small greek letters $\alpha, \beta, \ldots$. These conventions apply whenever these symbols do not obviously denote functions. Whenever ε

occurs in a statement, it is asserted that the statement is valid for any fixed positive value of ε. Note that if $A \ll P^\varepsilon$ and $B \ll P^\varepsilon$, then we may conclude that $AB \ll P^\varepsilon$. The leading parameter is P, and all statements are true whenever P exceeds a certain real number P_0 that depends only on k.

Vectors are in bold face, $\mathbf{x} = (x_1, \ldots, x_s)$, the dimension s will depend on the context. To avoid ambiguity, the greatest common factor of a and b is $(a; b)$.

The number of $\mathbf{u} \in \mathbb{N}^j$ with $u_1 u_2 \ldots u_j = n$ is denoted by $\mathrm{d}_j(n)$, and we write $\mathrm{d}(n) = \mathrm{d}_2(n)$ for the number of divisors of n. Similarly, Hooley's functions are defined by

$$\Delta_j(n) = \max_{\xi_1, \ldots, \xi_{j-1}} \#\{\mathbf{u} \in \mathbb{N}^j : u_1 u_2 \ldots u_j = n, \, \xi_i < \log u_i \le \xi_i + 1 \ (1 \le i < j)\},$$

and again we put $\Delta(n) = \Delta_2(n)$. Further, $\Omega(n)$ denotes the total number of prime factors of n, counted with multiplicity. Less standard, but common in related literature are the abbreviations

$$(1.7) \qquad L = \log P, \quad K = 2^{k-1}, \quad J = 2^{j-1}.$$

Acknowledgements. The authors thank the anonymous referee for his very detailed inspection of the manuscript. While this paper was refereed, further progress with the problems considered here was made in the case $k = 3$, see T.D. Wooley, Mean value estimates for odd cubic Weyl sums, arXiv:1401.7152.

2. The generating function

Central to the major arc analysis is a good approximation to $f(\alpha, \beta)$ when α, β are near rational numbers a/q, b/q. The approximating function is built from the expressions

$$(2.1) \qquad S(q, a, b) = \sum_{x=1}^{q} e\big((ax^k + bx)/q\big), \quad v(\xi, \zeta) = \int_0^P e(\xi t^k + \zeta t)\, \mathrm{d}t.$$

For frequent use later on, recall the estimates of Hua

$$(2.2) \qquad S(q, a, b) \ll q^{(k-1)/k+\varepsilon}, \quad v(\xi, \zeta) \ll P\big(1 + P^k|\xi| + P|\zeta|\big)^{-1/k}$$

that are valid for any real numbers ξ, ζ and natural numbers a, b, q subject to $(q; a; b) = 1$ (see Vaughan [29], Theorems 7.1 and 7.3).

Theorem 3. *Let* $q \in \mathbb{N}$, $a, b \in \mathbb{Z}$ *with* $(a; q) = 1$. *Let* $\alpha, \beta \in \mathbb{R}$ *and* $\alpha = \frac{a}{q} + \xi$, $\beta = \frac{b}{q} + \zeta$ *with* $|\zeta| \le 1/(2q)$. *Then*

$$(2.3) \qquad f(\alpha, \beta) = q^{-1} S(q, a, b) v(\xi, \zeta) + O\big(q^{(k-1)/k+\varepsilon}(1 + P^k|\xi|)^{1/2}\big).$$

If further $|\xi| \le 1/(4kqP^{k-1})$, *then*

$$(2.4) \qquad f(\alpha, \beta) = q^{-1} S(q, a, b) v(\xi, \zeta) + O\big(q^{(k-1)/k+\varepsilon}\big).$$

While we still work under the condition that $k \ge 3$, it may be worth pointing out that the conclusions in Theorem 3 and their proofs below remain valid when $k = 2$. However, when $k = 2$, a stronger version of Theorem 3 (with $\varepsilon = 0$) was obtained recently by Vaughan [30]. Apparently, Theorem 3 is new for all $k \ge 3$ and the best estimate available hitherto is the special case of [29], Theorem 7.2, which gives (2.3) with the error term inflated to $1 + P^k|\xi| + P|\zeta|$. It is vital for our later work that both q and $|\xi|$ occur in (2.3) with exponents below 1.

Our proof of Theorem 3 is an adaptation of a standard argument for the classical Weyl sum $f(\alpha, 0)$. We follow Vaughan [29], pp. 43–44, quite closely, but differences in detail justify a moderately detailed exposition. By (1.4), (2.1) and orthogonality of additive characters, one confirms the initial identity

$$(2.5) \qquad f(\alpha, \beta) = \frac{1}{q} \sum_{-q/2 < r \leq q/2} S(q, a, b - r) f\big(\xi, \zeta + (r/q)\big).$$

We apply a truncated Poisson summation formula to $f\big(\xi, \zeta + \frac{r}{q}\big)$. The phase $F(t) = \xi t^k + \big(\zeta + \frac{r}{q}\big)t$ has derivative $F'(t) = k\xi t^{k-1} + \zeta + \frac{r}{q}$, which is monotonic for $t \geq 0$, and for $0 \leq t \leq P$ and $|r| \leq \frac{1}{2}q$, one has $|F'(t)| \leq k|\xi|P^{k-1} + \frac{1}{2q} + \frac{1}{2}$. With $H = [k|\xi|P^{k-1}] + 2$, we have now verified the hypotheses of [29], Lemma 4.2, which gives

$$f\big(\xi, \zeta + \tfrac{r}{q}\big) = \sum_{|h| \leq H} v\Big(\xi, \zeta + \frac{r}{q} - h\Big) + O(\log H).$$

One uses this within (2.5). On writing $r - qh = m$ and $M = q(H + \frac{1}{2})$, this produces

$$f(\alpha, \beta) = \frac{1}{q} \sum_{-M < r \leq M} S(q, a, b - m) v\Big(\xi, \zeta + \frac{m}{q}\Big) + E$$

in which

$$E \ll \frac{\log H}{q} \sum_{|r| \leq q/2} |S(q, a, b - r)|.$$

Here, one isolates the term $m = 0$ and then applies (2.2) to all other terms to conclude that
$$(2.6)$$
$$f(\alpha, \beta) - q^{-1} S(q, a, b) v(\xi, \zeta) \ll q^{\varepsilon - 1/k} \sum_{1 \leq |m| \leq M} \big| v\big(\xi, \zeta + \tfrac{m}{q}\big)\big| + q^{(k-1)/k + \varepsilon} \log H.$$

We proceed to deduce (2.4). In the admissible range for ξ, one has $H = 2$. We take $F(t) = \xi t^k + \big(\zeta + \frac{m}{q}\big)t$ in Lemma 4.2 of Titchmarsh [25]. One has $F'(t) = \frac{m}{q} + R$ with $|R| \leq \frac{3}{4q}$ for $0 \leq t \leq P$, so that F' does not change sign and satisfies $|F'(t)| \geq \frac{|m|}{4q}$. The estimate provided by [25], Lemma 4.2, then shows that $v\big(\xi, \zeta + \frac{r}{q}\big) \ll q/|m|$, and (2.4) is immediate from (2.6).

It remains to prove (2.3) for $|\xi| \geq 1/(4kqP^{k-1})$, as we now assume. Let F be as before. Its derivative $F'(t) = k\xi t^{k-1} + \zeta + \frac{m}{q}$ is still monotonic for $t \geq 0$, but may have a zero in $[0, P]$. We therefore apply a stationary phase argument. For $1 \leq |m| \leq M$, let $\mathscr{T}(m)$ be the set of all $t \in [0, P]$ where $|F'(t)| \geq \frac{|m|}{4q}$. This is an interval or the union of two intervals, so that [25], Lemma 4.2, still shows that

$$(2.7) \qquad \int_{\mathscr{T}(m)} e(F(t)) \, dt \ll q/|m|.$$

It remains to estimate the contribution from $[0, P] \setminus \mathscr{T}(m)$. If this is non-empty, then $t \in [0, P] \setminus \mathscr{T}(m)$ satisfies

$$\Big| k\xi t^{k-1} + \zeta + \frac{m}{q} \Big| \leq \frac{|m|}{4q}.$$

This implies that

$$(2.8) \qquad \frac{3|m|}{4q} \leq |k\xi t^{k-1} + \zeta| \leq \frac{5|m|}{4q},$$

For $1 \leq |m| \leq M$, let

$$\delta = \delta(m) = |\xi|^{1/(2k-2)} \big(|m|/q\big)^{(k-2)/(2k-2)}.$$

We claim that

$$(2.9) \qquad \int_{[0,P]\setminus \mathscr{T}(m)} e(F(t))\, \mathrm{d}t \ll \delta^{-1}.$$

To see this, first note that an argument similar to the above shows that the set of all $0 \leq t \leq P$ with $\delta \leq |F'(t)| \leq \frac{|m|}{4q}$ contributes $O(\delta^{-1})$. For the remaining t, one has $|k\xi t^{k-1} + \zeta + \frac{m}{q}| \leq \delta$, and if t_1, t_2 satisfy this inequality, one finds that $k|\xi||t_1^{k-1} - t_2^{k-1}| \leq 2\delta$. But (2.8) together with $|\zeta| \leq 1/(2q)$ implies that $t_j^{k-1} \geq \frac{|m|}{4k|\xi|q}$, so that the binomial expansion yields

$$|t_1 - t_2| \leq \frac{2\delta}{k|\xi|} \Big(\frac{4k|\xi|q}{|m|}\Big)^{(k-2)/(k-1)} \ll \frac{1}{\delta}.$$

Hence, the set of $t \in [0,P]$ with $|F'(t)| \leq \delta$ has measure $O(1/\delta)$, and (2.9) follows. On combining (2.7) and (2.9) to an estimate for $v\big(\xi, \zeta + \frac{m}{q}\big)$, we find that

$$\sum_{1 \leq |m| \leq M} \big|v\big(\xi, \zeta + \tfrac{m}{q}\big)\big| \ll q \sum_{m \leq M} \frac{1}{m} + \sum_{m \leq 1+2kq|\xi|P^{k-1}} \frac{1}{\delta(m)}$$
$$\ll q^{1+\varepsilon}(1 + P^k|\xi|)^{1/2}.$$

In view of (2.6), the desired bound (2.3) is immediate. $\square$

We now apply Theorem 3 to establish a strong from of Weyl's inequality for $f(\alpha, \beta)$. Recall (1.4), (1.7) and the definition of $\mathfrak{m}$ in Theorem 2.

Lemma 1. *Uniformly in $\beta \in \mathbb{R}$, $\alpha \in \mathfrak{m}$, one has*
$$|f(\alpha, \beta)|^K \ll P^{K-1}L^{1+\varepsilon}.$$

The proof is essentially that of Lemma 1 in Vaughan [26], suitably generalised to $k \geq 3$. However, our analysis relies on Theorem 3 rather than [29, Theorem 4.1], and some care is required to accomodate the weaker error estimates. In particular, it turns out that mimicry of the argument outlined on p. 131 of [26] leads to a satisfactory bound only in the case when $k \geq 4$, which we temporarily suppose from now on.

Let $\delta = 1/(100k)$ and $\alpha \in \mathfrak{m}$. By Dirichlet's theorem, there are coprime a, q with $1 \leq q \leq P^{k-1-\delta}$ and $|q\alpha - a| \leq P^{1+\delta-k}$. First suppose that $q \leq P^\mu$ where $\mu = 1 + \frac{1}{k+3}$. One checks that $\frac{k-1}{k}\mu < 1 - \frac{1}{K}$ holds for all $k \geq 4$. Hence by Theorem 3 and (2.2),

$$(2.10) \qquad f(\alpha, \beta) \ll q^{\varepsilon - 1/k} P\big(1 + P^k|\alpha - \tfrac{a}{q}|\big)^{-1/k} + P^{1-1/K}.$$

Since $\alpha \in \mathfrak{m}$, we have $q > Q$ or $|\alpha - a/q| \geq Q/(qP^k)$, and in both cases it follows that

$$(2.11) \qquad f(\alpha, \beta) \ll Q^{\varepsilon - 1/k} P + P^{1-1/K} \ll P^{1-1/K},$$

as required.

This leaves the case where $q > P^\mu$. Then $|\alpha - a/q| \le P^{1+\delta-\mu-k}$, and one checks that $1 + \delta - \mu = -\frac{1}{k+3} + \delta \le -\frac{1}{K}$. Consequently,

$$f(\alpha, \beta) - f\left(\frac{a}{q}, \beta\right) = \sum_{x \le P} e(\beta x)\left(e(\alpha x^k) - e\left(\frac{a}{q} x^k\right)\right) \ll P^{1-1/K}.$$

Thus, we are reduced to estimating $f(a/q, \beta)$. We begin by observing that the substitution $y = x + h$ produces

$$|f(\alpha, \beta)|^2 = \sum_{x,y \le P} e\left(\alpha(y^k - x^k) + \beta(y - x)\right)$$

(2.12)
$$= \sum_{|h| < P} e(\beta h) \sum_{\substack{1 \le x \le P \\ 1 \le x + h \le P}} e\left(\alpha((x+h)^k - x^k)\right)$$

$$\le P + \sum_{0 < |h| < P} \left| \sum_{\substack{1 \le x \le P \\ 1 \le x + h \le P}} e\left(\alpha((x+h)^k - x^k)\right) \right|.$$

Note that β is absent in this inequality. We now take $\alpha = a/q$ and repeat Weyl differencing in the usual way. Then, as in [26], p. 131, one arrives at

$$|f(a/q, \beta)|^K \ll P^{K-1} + P^{K-k} \sum_{\substack{1 \le h_j \le P \\ 1 \le j \le k-1}} \min\left(P, \left\| \frac{k! a h_1 \ldots h_{k-1}}{q} \right\|^{-1}\right),$$

and one may then complete the estimation in the same way as in the final part of the proof of [26, Lemma 1], but using Hall and Tenenbaum [13, Theorem 70] for $\Delta_{k-1}(n)$ instead of referencing Hooley [16]. This completes the proof of Lemma 1 when $k \ge 4$.

Now suppose that $k = 3$. More care is required in the part that relies on Weyl differencing. On applying Cauchy's inequality to

$$f(\alpha, \beta) = \sum_{u=1}^{2} \sum_{\substack{x \le P \\ x \equiv u \bmod 2}} e(\alpha x^3 + \beta x)$$

one finds that

$$|f(\alpha, \beta)|^2 \le 2 \sum_{\substack{x,y \le P \\ x \equiv y \bmod 2}} e\left(\alpha(x^k - y^k) + \beta(x - y)\right),$$

and the substitution $2z = x + y$, $2h = x - y$ transforms this to

$$|f(\alpha, \beta)|^2 \le 2 \sum_{|h| \le P/2} \sum_{z \in I(h)} e(2\beta h) e\left(2\alpha h(3z^2 + h^2)\right)$$

where $I(h)$ is the subinterval of $[1, P]$ described by the inequalities $1 \le z + h \le P$, $1 \le z - h \le P$. It follows that

$$|f(\alpha, \beta)|^2 \le 2 \sum_{|h| \le P/2} \left| \sum_{z \in I(h)} e\left(6\alpha h z^2\right) \right|$$

$$\le 2P + 4 \sum_{1 \le h \le P/2} \left| \sum_{z \in I(h)} e\left(6\alpha h z^2\right) \right|.$$

By Cauchy's inequality,

$$|f(\alpha,\beta)|^4 \ll P^2 + P \sum_{0<|h|\leq P/2} \Big| \sum_{z\in I(h)} e\big(6\alpha hz^2\big)\Big|^2.$$

The double sum over h,z here is that same as the one estimated in Lemma 4 of Vaughan [26] with $H = P/2$, except that in that lemma, 6α is 3α, and the range for z is $1 \leq z \leq P$. An inspection of the proof shows that these changes are irrelevant, and that the conclusion is still valid in our context. Moreover, similar to an earlier comment, the use of Theorem 70 of Hall and Tenenbaum [13] within the proof of [26, Lemma 4] reduces the factor $L^{4/\pi+\varepsilon}$ in that lemma to $L^{1+\varepsilon}$. Hence, on choosing $1 \leq q \leq P^{2-\delta}$ and $a \in \mathbb{Z}$ with $|q\alpha-a| \leq P^{\delta-2}$, the augmented form of [26, Lemma 4] now shows that $|f(\alpha,\beta)|^4 \ll P^3 L^{1+\varepsilon}$ holds in all cases where $q > P^{1+\delta}$. When $q \leq P^{1+\delta}$, then Theorem 3 and (2.2) yield (2.10) for $k = 3$, and one then also confirms (2.11) for $k = 3$. The proof of Lemma 1 is complete. $\qquad\square$

We close this section with a technical observation concerning the exponential sum defined in (2.1).

Lemma 2. *Let p be a prime, and suppose that a, b are integers with $p \mid a$ and $p \nmid b$. Then, for all $l \in \mathbb{N}$, one has $S(p^l, a, b) = 0$.*

Proof. On substituting $x = p^{l-1}y + z$ in (2.1), one finds that

$$S(p^l, a, b) = \sum_{x=1}^{p^l} e\Big(\frac{ax^k}{p^l} + \frac{bx}{p^l}\Big) = \sum_{y=1}^{p} e\Big(\frac{by}{p}\Big) \sum_{z=1}^{p^{l-1}} e\Big(\frac{az^k + bz}{p^l}\Big) = 0.$$

3. Preparatory mean value estimates

We begin with certain divisor sums that are routinely estimated by van der Corput's method. Let $t \in \mathbb{N}$, put $\mathbf{x} = (x_1,\ldots,x_t)$ and define the sums

$$(3.1)\qquad \Lambda(\mathbf{x},\mathbf{y}) = \sum_{j=1}^{t}(x_j - y_j), \quad \mathrm{M}(\mathbf{x},\mathbf{y}) = \sum_{j=1}^{t}(x_j^k - y_j^k).$$

Lemma 3. *Let $t \in \mathbb{N}$, $l \in \mathbb{N}$. Then, there exists positive numbers δ and η such that*

(3.2)
$$\sum_{\substack{x_j\leq P,\, y_j\leq P \\ \mathrm{M}(\mathbf{x},\mathbf{y})\neq 0}} \mathrm{d}\big(\mathrm{M}(\mathbf{x},\mathbf{y})\big)^l \ll P^{2t}L^{\eta}, \qquad \sum_{\substack{x_j\leq P,\, y_j\leq P \\ \mathrm{M}(\mathbf{x},\mathbf{y})\neq 0}} \mathrm{d}\big(\mathrm{M}(\mathbf{x},\mathbf{y})\big)^l e^{\delta\Omega(\mathrm{M}(\mathbf{x},\mathbf{y}))} \ll P^{2t}L^{\eta},$$

(3.3)
$$\sum_{\substack{x_j\leq P,\, y_j\leq P \\ \mathrm{M}(\mathbf{x},\mathbf{y})\neq 0 \\ \Lambda(\mathbf{x},\mathbf{y})=0}} \mathrm{d}\big(\mathrm{M}(\mathbf{x},\mathbf{y})\big)^l \ll P^{2t-1}L^{\eta}, \qquad \sum_{\substack{x_j\leq P,\, y_j\leq P \\ \mathrm{M}(\mathbf{x},\mathbf{y})\neq 0 \\ \Lambda(\mathbf{x},\mathbf{y})=0}} \mathrm{d}\big(\mathrm{M}(\mathbf{x},\mathbf{y})\big)^l e^{\delta\Omega(\mathrm{M}(\mathbf{x},\mathbf{y}))} \ll P^{2t-1}L^{\eta}.$$

Proof. The first estimate in (3.2) is a special case of Theorem 3 of Hua [22], and the second can be established by a development of the underlying method (see also Lemma 1 of Vaughan [27]). To verify the first bound in (3.3), substitute $y_t = \sum_{j=1}^{t-1}(x_j - y_j) + x_t$ into $M(\mathbf{x},\mathbf{y})$. Then the sum on the left hand side of (3.3)

does not exceed the sum

$$\sum_{x_1,\ldots,x_t \le P} \sum_{y_1,\ldots,y_{t-1} \le P} \mathrm{d}\Big(\sum_{j=1}^{t-1}(x_j^k - y_j^k) + x_t^k - \Big(\sum_{j=1}^{t-1}(x_j - y_j) + x_t\Big)^k\Big)^l,$$

and the desired bound again follows from Theorem 3 of Hua [22]. Finally, once again a development of the method also yields the second estimate in (3.3). □

We now consider the pair of diophantine equations

$$(3.4) \qquad x_1^k + x_2^k + x_3^k = y_1^k + y_2^k + y_3^k, \quad x_1 + x_2 + x_3 = y_1 + y_2 + y_3.$$

Let $\Gamma(P)$ denote the number of solutions of (3.4) with $1 \le x_j \le P$, $1 \le y_j \le P$ such that x_1, x_2, x_3 is not a permutation of y_1, y_2, y_3. Then, by Vaughan and Wooley [31] when $k = 3$, and by Greaves [11] when $k \ge 4$, there exists $\gamma > 0$ such that $\Gamma(P) \ll P^{3-\gamma}$. With this value of γ, this immediately implies the following estimate.

Lemma 4. *For $\mathscr{U} \subset \{1, 2, \ldots, [P]\}$, let $U(P)$ denote the number of solutions of (3.4) with $x_j \in \mathscr{U}$, $y_j \in \mathscr{U}$. Then*

$$U(P) \ll \big(\#\mathscr{U}\big)^3 + P^{3-\gamma}.$$

The next lemma should be compared with the classical lemma of Hua in the theory of Waring's problem.

Lemma 5. *There is a positive number η such that whenever $2 \le j \le k$, then*

$$\int_0^1 \int_0^1 |f(\alpha, \beta)|^{2^j+2}\, \mathrm{d}\alpha\, \mathrm{d}\beta \ll P^{2^j-j+1} L^\eta.$$

Proof. Let $j = 2$. By orthogonality, the integral equals $U(P)$ with $\mathscr{U} = [1, P] \cap \mathbb{Z}$ so that this case of Lemma 5 is a consequence of Lemma 4.

We now suppose that the conclusion of Lemma 5 has been established for a particular value of j with $2 \le j < k$, and proceed by induction. Recall that $J = 2^{j-1}$, and return to (2.12). Repeated Weyl differencing via Cauchy's inequality then gives

$$(3.5) \qquad |f(\alpha, \beta)|^{2J} \ll P^{2J-j-1} \sum_{|h_1| < P} \cdots \sum_{|h_j| < P} \sum_{x \in I(\mathbf{h})} e\big(\alpha h_1 \ldots h_j Q_{\mathbf{h}}(x)\big)$$

where $I(\mathbf{h}) \subset [1, P]$ is a suitable interval and $Q_{\mathbf{h}} \in \mathbb{Z}[X]$ has degree $k - j$. Now let $r(l)$ denote the number of solutions of

$$(3.6) \qquad\qquad\qquad h_1 \ldots h_j Q_{\mathbf{h}}(x) = l$$

with all variables $h_1, \ldots, h_j, x$ subject to the summation conditions in the preceding display. Then

$$|f(\alpha, \beta)|^{2J} \ll P^{2J-j-1} \sum_{l \in \mathbb{Z}} r(l) e(-\alpha l).$$

By (1.4),

$$|f(\alpha, \beta)|^{2J+2} = \sum_{\substack{x_i \le P,\, y_i \le P \\ 1 \le i \le J+1}} e\big(\alpha \mathrm{M}(\mathbf{x}, \mathbf{y}) + \beta \Lambda(\mathbf{x}, \mathbf{y})\big)$$

where Λ and M are the forms defined in (3.1) with $t = J + 1$. Hence, by orthogonality,

$$(3.7) \qquad \int_0^1 \int_0^1 |f(\alpha, \beta)|^{4J+2} \, \mathrm{d}\alpha \, \mathrm{d}\beta \ll P^{2J-j-1} \sum_{\substack{\mathbf{x}, \mathbf{y} \\ \Lambda(\mathbf{x}, \mathbf{y}) = 0}} r\big(\mathrm{M}(\mathbf{x}, \mathbf{y})\big).$$

One has $r(0) \ll P^j$, and $r(l) \ll \mathrm{d}(|l|)^{j+1}$ for $l \neq 0$. Hence the contribution to (3.7) of all terms where $\mathrm{M}(\mathbf{x}, \mathbf{y}) \neq 0$ can be estimated by (3.3), and does not exceed $P^{4J-j} L^\eta$. This leaves the solutions of $\mathrm{M}(\mathbf{x}, \mathbf{y}) = \Lambda(\mathbf{x}, \mathbf{y}) = 0$, and by orthogonality again, these contribute to (3.7) at most

$$P^{2J-1} \int_0^1 \int_0^1 |f(\alpha, \beta)|^{2J+2} \, \mathrm{d}\alpha \, \mathrm{d}\beta \ll P^{4J-j} L^\eta.$$

This completes the induction step, and the proof of the lemma. $\qquad \square$

Let κ denote a positive number sufficiently large in terms of k, and write

$$(3.8) \qquad\qquad D = L^\kappa, \quad D' = P^{1/(6k)}.$$

Let $\mathscr{E}$ denote the set of all $n \leq P$ with no prime factor in the interval $[D, D']$. Then, uniformly for $d \leq P$ and $m \in \mathbb{N}$, the number $E_{d,m}(P)$ of $n \in \mathscr{E}$ with $n \equiv m \bmod d$ satisfies

$$(3.9) \qquad\qquad E_{d,m}(P) \ll \frac{P \log \log P}{\varphi(d) \log(2P/D)}$$

(see Lemma 2 of Vaughan [27] or Halberstam-Richert [12], Theorem 3.4). We now define the exponential sum

$$(3.10) \qquad\qquad h(\alpha, \beta) = \sum_{x \in \mathscr{E}} e(\alpha x^k + \beta x).$$

The next lemma is a considerable refinement of the preceding lemma, and crucial for all later work.

Lemma 6. *Let* $2 \leq j \leq k$. *Then*

$$\int_0^1 \int_0^1 |h(\alpha, \beta)|^{2^j + 2} \, \mathrm{d}\alpha \, \mathrm{d}\beta \ll P^{2^j - j + 1} L^{\varepsilon - 1 - \frac{1}{2} j(j-1)}.$$

Proof. When $j = 2$, this follows from orthogonality, Lemma 4 and (3.9). Now suppose the estimate is known for a particular value of j with $2 \leq j < k$, and proceed by induction. The argument to follow is very similar to that on pp. 14–19 of [27], so we shall be brief whenever the modifications to [27] are evident.

We begin by applying Weyl differencing to (3.10). The first differencing is performed as in (2.12), and delivers the initial inequality

$$|h(\alpha, \beta)|^2 \leq \#\mathscr{E} + \sum_{0 < |h| < P} \Big| \sum_{\substack{x \in \mathscr{E} \\ x + h \in \mathscr{E}}} e\big(\alpha((x+h)^k - x^k)\big) \Big|.$$

Since $j \geq 2$, we have to difference further to reach the inequality

$$|h(\alpha, \beta)|^{2^j} \ll P^{2^j - j - 1} \big(\#\mathscr{E}\big)^j + P^{2^j - j - 1} \sum_{\substack{1 \leq |h_i| < P \\ 1 \leq i \leq j}} \sum_x e\big(\alpha h_1 \ldots h_j Q_j(x, \mathbf{h})\big)$$

in which the sum over x is subject to the constraints $x \in \mathscr{E}$ and $x + h_i \in \mathscr{E}$ $(1 \le i \le j)$, and Q_j is as in (3.5). Now multiply with $|h(\alpha, \beta)|^{2^j + 2}$ and integrate. Then, recalling that $J = 2^{j-1}$, one finds that

(3.11)

$$\int_0^1 \int_0^1 |h(\alpha, \beta)|^{4J+2} \, \mathrm{d}\alpha \, \mathrm{d}\beta$$
$$\ll P^{2J-j-1} \big(\#\mathscr{E}\big)^j \int_0^1 \int_0^1 |h(\alpha, \beta)|^{2J+2} \, \mathrm{d}\alpha \, \mathrm{d}\beta + P^{2J-j-1} \sum_{\substack{\mathbf{x}, \mathbf{y} \\ \Lambda(\mathbf{x}, \mathbf{y}) = 0}} r\big(\mathrm{M}(\mathbf{x}, \mathbf{y})\big)$$

where $r(l)$ has the same meaning as in (3.6), M and Λ are defined by (3.1) with $t = J + 1$, and $\mathbf{x}, \mathbf{y}$ are subject to $x_j, y_j \in \mathscr{E}$, $\mathrm{M}(\mathbf{x}, \mathbf{y}) \ne 0$. By the induction hypothesis, the first term on the right hand side does not exceed

$$\ll P^{2J-1} L^{\varepsilon - j} P^{2J-j+1} L^{\varepsilon - 1 - \frac{1}{2}j(j-1)} \ll P^{4J-j} L^{\varepsilon - 1 - \frac{1}{2}j(j+1)},$$

as required. In view of (3.11), the induction will be complete once the inequality

(3.12) $$\sum_{\substack{\mathbf{x}, \mathbf{y} \\ \Lambda(\mathbf{x}, \mathbf{y}) = 0}} r\big(|\mathrm{M}(\mathbf{x}, \mathbf{y})|\big) \ll P^{2J+1} L^{\varepsilon - 1 - \frac{1}{2}j(j+1)}$$

is established; here and later the sum is subject to same conditions as in (3.11). Let $D = L^\kappa$ be as in (3.8). The contribution to (3.12) arising from summands with $|\mathrm{M}(\mathbf{x}, \mathbf{y})| \le (P/D)^k$ is small. To see this, we use Cauchy's inequality to infer that

(3.13) $$\sum_{\substack{\Lambda(\mathbf{x}, \mathbf{y}) = 0 \\ 1 \le |\mathrm{M}(\mathbf{x}, \mathbf{y})| \le (P/D)^k}} r(|\mathrm{M}(\mathbf{x}, \mathbf{y})|) \le \big(T_1 T_2\big)^{1/2}$$

where

$$T_1 = \sum_{\Lambda(\mathbf{x}, \mathbf{y}) = 0} \mathrm{d}(|\mathrm{M}(\mathbf{x}, \mathbf{y})|)^{2j}$$

and T_2 denotes the number of all $1 \le x_i, y_i \le P$, $1 \le i \le J + 1$ with

(3.14) $$\Lambda(\mathbf{x}, \mathbf{y}) = 0, \quad 1 \le |\mathrm{M}(\mathbf{x}, \mathbf{y})| \le (P/D)^k.$$

By (3.3), we have $T_1 \ll P^{2J+1} L^\eta$. With any solution $\mathbf{x}, \mathbf{y}$ counted by T_2, we associate the numbers

$$u = \sum_{i=2}^{J+1} (y_i - x_i), \quad w = \sum_{i=2}^{J+1} (y_i^k - x_i^k).$$

Let T_2' denote the number of solutions counted by T_2 where $0 \le u \le P/D$, and let T_2'' denote the number of those solutions counted by T_2 where $P/D < u \le JP$. Then, by symmetry in $\mathbf{x}$ and $\mathbf{y}$, it follows that

$$T_2 \le 2(T_2' + T_2'').$$

To estimate T_2', we consider one particular choice of $x_2, \ldots, x_{J+1}, y_3, \ldots, y_{J+1}$; there are $O(P^{2J-1})$ possibilities for this. The conditions that $0 \le u \le P/D$ leaves $O(P/D)$ choices for y_2. Once these variables are fixed, u is also fixed, and so $x_1 - y_1 = u$ leaves $O(P)$ choices for the pair x_1, y_1. This shows that $T_2' \ll P^{2J+1}/D$.

The initial treatment of T_2'' is similar. Fix one of the $O(P^{2J})$ choices for $x_2, \ldots, x_{J+1}, y_2, \ldots, y_{J+1}$ with $u > P/D$. This fixes u and w, and by (3.14) it

remains to count the x_1, y_1 with $x_1 - y_1 = u$, $|x_1^k - y_1^k - w| \leq (P/D)^k$. We eliminate x_1, and consider the inequality

$$(3.15) \qquad |(y_1 + u)^k - y_1^k - w| \leq (P/D)^k.$$

Let z_1 and z_2 be two solutions (for y_1) of (3.15) with $1 \leq z_1 \leq z_2 \leq P$. Then

$$(3.16) \qquad \left|(z_2 + u)^k - z_2^k - (z_1 + u)^k + z_1^k\right| \leq 2(P/D)^k.$$

A direct computation yields

$$(z_2 + u)^k - z_2^k - (z_1 + u)^k + z_1^k = k(k-1) \int_0^u \int_0^{z_2 - z_1} (z_1 + \zeta + \xi)^{k-2} \, d\zeta \, d\xi$$

$$\geq k(k-1) \int_0^u \int_0^{z_2 - z_1} \zeta^{k-2} \, d\zeta \, d\xi = ku(z_2 - z_1)^{k-1}.$$

Recalling that one has $u > P/D$ in the current context, one infers from (3.16) that $z_2 - z_1 \leq P/D$. This shows that (3.15) has at most $2(P/D) + 1$ solutions in y_1. Collecting together, we see that $T_2'' \ll P^{2J+1}/D$, that $T_2 \ll P^{2J+1}/D$, and by (3.13),

$$(3.17) \qquad \sum_{\substack{\Lambda(\mathbf{x},\mathbf{y})=0 \\ |\mathrm{M}(\mathbf{x},\mathbf{y})| \leq (P/D)^k}} r(|\mathrm{M}(\mathbf{x},\mathbf{y})|) \ll P^{2J+1} L^{\frac{1}{2}(\eta - \kappa)}.$$

We are reduced to estimating

$$(3.18) \qquad \Upsilon = \sum_{\substack{\Lambda(\mathbf{x},\mathbf{y})=0 \\ |\mathrm{M}(\mathbf{x},\mathbf{y})| > (P/D)^k}} r(|\mathrm{M}(\mathbf{x},\mathbf{y})|).$$

Let $\varrho(n)$ denote the number of solutions of the system $|\mathrm{M}(\mathbf{x},\mathbf{y})| = n$, $\Lambda(\mathbf{x},\mathbf{y}) = 0$ with $x_i, y_i \in \mathscr{E}$ $(1 \leq i \leq J+1)$. Then $\varrho(n) = 0$ for $n > (J+1)P^k$, and the definition of $r(n)$ shows that for $(P/D)^k < n \leq 2JP^k$ one has $r(n) \ll (\log \log P)^j \Delta_{j+1}(n)$. It follows that

$$(3.19) \qquad \Upsilon \ll L^\varepsilon \sum_n \Delta_{j+1}(n) \varrho(n).$$

Let ν be a (small) positive number, and let δ be as in Lemma 3. Let

$$Y = \exp\left(\frac{\delta \nu L}{(\eta + k^2) \log L}\right).$$

Let

$$\mathscr{M} = \{n \in \mathbb{N} \colon p \mid n \Rightarrow p \leq Y\}, \quad \mathscr{N} = \{n \in \mathbb{N} \colon p \mid n \Rightarrow p > Y\}.$$

Then, any $n \in \mathbb{N}$ has a unique factorisation $n = n^* n^\dagger$ with $n^* \in \mathscr{M}$, $n^\dagger \in \mathscr{N}$. Note that $n^* > P^\nu$ implies $\Omega(n) \log Y \geq \log n^* > \nu L$ so that $\delta \Omega(n) > (\eta + k^2) \log L$. By (3.3),

$$(3.20) \qquad \sum_{n^* > P^\nu} \Delta_{j+1}(n) \varrho(n) \ll \sum_n \mathrm{d}(n)^{2j+2} e^{\delta \Omega(n)} \varrho(n) L^{-\eta - k^2} \ll P^{2J+1} L^{-k^2}$$

which is acceptable. For the complementary portion of (3.19), we have

$$\sum_{n^* \leq P^\nu} \Delta_{j+1}(n)\varrho(n) \leq \sum_{\substack{m \in \mathcal{M} \\ m \leq P^\nu}} \sum_{n \in \mathcal{N}} \Delta_{j+1}(mn)\varrho(mn)$$

$$\leq \sum_{\substack{m \in \mathcal{M} \\ m \leq P^\nu}} \Delta_{j+1}(m) \sum_{n \in \mathcal{N}} \mathrm{d}_{j+1}(n)\varrho(mn).$$

We now require the following simple observation.

Lemma 7. *Let $k \geq 3$ and $\nu > 0$. There exists a number L_0 depending only on k and ν with the property that for any $2 \leq j \leq k-1$ and any $n < 2JP^k$ there is a divisor n_1 of n with $n_1 \leq P^\nu$ and $\mathrm{d}_{j+1}(n) \ll j^{L_0 \Omega(n_1)}$.*

Although not highlighted as a lemma, the conclusion of Lemma 7 is established *inter alia* in [27], starting on p. 16 after [27, (3.15)]. □

We apply Lemma 7 to the inner sum in the previous display and obtain

$$(3.21) \quad \sum_{n^* \leq P^\nu} \Delta_{j+1}(n)\varrho(n) \ll \sum_{\substack{m \in \mathcal{M} \\ m \leq P^\nu}} \Delta_{j+1}(m) \sum_{\substack{n_1 \in \mathcal{N} \\ n_1 \leq P^\nu}} j^{L_0 \Omega(n_1)} \sum_{n \equiv 0 \bmod mn_1} \varrho(n).$$

It will be convenient to write $d = mn_1$. Then, the conditions active in (3.21) imply that $d \leq P^{2\nu} \leq \sqrt{P}$. Further, the sum $\sum_{n \equiv 0 \bmod d} \varrho(n)$ does not exceed the number of solutions of

$$(3.22) \qquad \sum_{i=1}^{J+1}(x_i - y_i) = 0, \quad \sum_{i=1}^{J+1}(x_i^k - y_i^k) \equiv 0 \bmod d$$

with $x_i, y_i \in \mathscr{E}$ $(1 \leq i \leq J+1)$. Let $\mathbf{a}, \mathbf{b}$ be a solution of the pair of congruences

$$(3.23) \qquad \sum_{i=1}^{J+1}(a_i - b_i) \equiv \sum_{i=1}^{J+1}(a_i^k - b_i^k) \equiv 0 \bmod d$$

and choose $x_i, y_i \in \mathscr{E}$ with $x_i \equiv a_i \bmod d$ for $1 \leq i \leq J+1$ and $y_i \equiv b_i \bmod d$ for $1 \leq i \leq J$. Then determine y_{J+1} through the linear equation in (3.22). By (3.23), it follows that $y_{J+1} \equiv b_{J+1} \bmod d$, and by (3.9), we infer that the number of solutions to (3.22) with $x_i, y_i \in \mathscr{E}$ and $x_i \equiv a_i \bmod d$, $y_i \equiv b_i \bmod d$ $(1 \leq i \leq J+1)$ does not exceed $O\big((P/d)^{2J+1}L^{\varepsilon-1-2J}\big)$. We conclude that

$$\sum_{n: \, d|n} \varrho(n) \ll P^{2J+1}L^{\varepsilon-1-2J}d^{-1-2J}S(d)$$

where $S(d)$ is the number of incongruent solutions to the pair of congruences (3.23). Since $S(d)$ is multiplicative, we deduce from (3.21) that

$$(3.24) \qquad \sum_{n^* \leq P^\nu} \Delta_{j+1}(n)\varrho(n) \ll P^{2J+1}L^{\varepsilon-1-2J}\Xi_1\Xi_2$$

in which

$$\Xi_1 = \sum_{\substack{m \in \mathcal{M} \\ m \leq P^\nu}} \frac{\Delta_{j+1}(m)S(m)}{m^{1+2J}}, \quad \Xi_2 = \sum_{\substack{n \in \mathcal{N} \\ n \leq P^\nu}} \frac{j^{L_0 \Omega(n)}S(n)}{n^{1+2J}}.$$

Further progress depends on upper bounds for $S(d)$ that we now derive. By (2.1) and orthogonality,

$$(3.25) \qquad S(d) = d^{-2} \sum_{a=1}^{d} \sum_{b=1}^{d} |S(q,a,b)|^{2J+2}.$$

But $j \geq 2$, so that $2J + 2 \geq 6$, and then

$$S(d) \leq d^{2J-6} \sum_{a=1}^{d} \sum_{b=1}^{d} |S(q,a,b)|^6 = d^{2J-4} S_0(d)$$

where $S_0(d)$ is the number of solutions of the congruences

$$\sum_{i=1}^{3} (u_i - v_i) \equiv \sum_{i=1}^{3} (u_i^k - v_i^k) \equiv 0 \bmod d.$$

In particular, $S_0(d)$ is a multiplicative function. We now have

$$\Xi_1 \leq \sum_{\substack{m \in \mathscr{M} \\ m \leq P^\nu}} \frac{\Delta_{j+1}(m) S_0(m)}{m^5}, \quad \Xi_2 \leq \sum_{\substack{n \in \mathscr{N} \\ n \leq P^\nu}} \frac{j^{L_0 \Omega(n)} S_0(n)}{n^5}.$$

By (2.1), whenever $d = (q; a; b)$, one has $S(q,a,b) = dS(q/d, a/d, b/d)$. Hence, on noting that $S_0(d)$ equals the right hand side of (3.25) with $J = 2$, and then collecting terms according to $(q; a; b)$, one readily confirms that

$$(3.26) \qquad \frac{S_0(m)}{m^4} = \sum_{q|m} A(q)$$

where

$$(3.27) \qquad A(q) = q^{-6} \sum_{\substack{a,b=1 \\ (a;b;q)=1}}^{q} |S(q,a,b)|^6.$$

By (3.26) and Möbius inversion, $A(q)$ is multiplicative. We shall prove momentarily that the series

$$(3.28) \qquad \sum_{q=1}^{\infty} \frac{A(q) \mathrm{d}_{j+1}(q)}{q}, \qquad \sum_{q \in \mathscr{N}} j^{L_0 \Omega(q)} \frac{A(q)}{q}$$

converge, and that the second series (which depends on P) is bounded above by a constant depending only on k. Once this is established, the proof of Lemma 6 is swiftly completed. Indeed, the familiar inequality $\Delta_{j+1}(uw) \leq \Delta_{j+1}(u) \mathrm{d}_{j+1}(w)$ now implies that

$$\Xi_1 \leq \sum_{\substack{m \in \mathscr{M} \\ m \leq P^\nu}} \frac{\Delta_{j+1}(m)}{m} \sum_{q|m} A(q) \leq \sum_{u \leq P^\nu} \frac{\Delta_{j+1}(u)}{u} \sum_{q \leq P^\nu} \frac{A(q) \mathrm{d}_{j+1}(q)}{q},$$

and by (3.28) the sum over q is bounded. Also, by Hall and Tenenbaum [13], Theorem 70, the sum over u is $O(L^{1+\varepsilon})$, whence $\Xi_1 \ll L^{1+\varepsilon}$. Similarly,

$$\Xi_2 \leq \sum_{\substack{n \in \mathscr{N} \\ n \leq P^\nu}} \frac{j^{L_0 \Omega(n)}}{n} \sum_{q|n} A(q) \leq \sum_{\substack{u \in \mathscr{N} \\ u \leq P^\nu}} \frac{j^{L_0 \Omega(u)}}{u} \sum_{\substack{q \in \mathscr{N} \\ q \leq P^\nu}} j^{L_0 \Omega(q)} \frac{A(q)}{q}.$$

Here again, the sum over q is bounded and

$$\sum_{\substack{u \in \mathscr{N} \\ u \leq P^{\nu}}} \frac{j^{L_0 \Omega(u)}}{u} = \prod_{Y < p \leq P^{\nu}} \left(1 + \sum_{l=1}^{\infty} \frac{j^{L_0 l}}{p^l}\right) \ll (\log \log P)^{j^{L_0}},$$

so that $\Xi_2 \ll L^{\varepsilon}$. On collecting together, it follows that the expression on the left hand side of (3.24) is $O\left(P^{2J+1} L^{\varepsilon - 2J}\right)$. But $2J \geq 1 + \frac{1}{2}j(j+1)$ for $j \geq 2$, and therefore, by (3.24), (3.19) and (3.20), the sums (3.18) and (3.17) (with $\kappa = \eta + k^2$) are both sufficiently small to imply (3.12). This completes the proof of Lemma 6.

In preparation for the discussion of the series in (3.28), we require an upper bound for $A(q)$. By multiplicativity, it will suffice to consider the case where $q = p^l$ is a prime power. In this case, one infers from (3.27) and Lemma 2 that

$$(3.29) \qquad A(p^l) = p^{-6l} \sum_{\substack{a=1 \\ p \nmid a}}^{p^l} \sum_{b=1}^{p^l} |S(p^l, a, b)|^6.$$

For a crude bound, note that whenever $p \nmid a$ one has $p^{-l} S(p^l, a, b) \ll p^{-l/k}$. This is slighly stronger than (2.1) but follows from the proof of [29], Theorem 7.1, or Chalk [8], for example. By (3.29), it follows that

$$A(p^l) \ll p^{-2l - 4l/k} \sum_{a=1}^{p^l} \sum_{b=1}^{p^l} |S(p^l, a, b)|^2,$$

and by orthogonality, one derives the estimate

$$(3.30) \qquad p^{-l} A(p^l) \ll p^{-4l/k}.$$

We proceed to establish the alternative estimate

$$(3.31) \qquad p^{-l} A(p^l) \ll p^{-2}$$

that is valid for all primes p with $p \nmid k$. Indeed, by Lemma 4.3 of Vaughan [29], one has $S(p, a, 0) \ll p^{1/2}$ whenever $p \nmid a$, and when $p \nmid b$, Lemma 4.1 of Vaughan [29] gives $S(p, a, b) \ll p^{1/2}$. By (3.29), it follows that $A(p) \ll p^{-1}$. This already confirms (3.31) when $l = 1$, and by (3.26), one also finds that

$$(3.32) \qquad S_0(p) \ll p^4.$$

Now let $l \geq 2$, and consider a solution of the system of congruences

$$(3.33) \qquad \sum_{i=1}^{3}(u_i - v_i) \equiv \sum_{i=1}^{3}(u_i^k - v_i^k) \equiv 0 \bmod p^l$$

with $1 \leq u_i, v_i \leq p^l$. Such a solution is said to be non-singular modulo p if the array

$$\begin{pmatrix} ku_1^{k-1} & ku_2^{k-1} & ku_3^{k-1} & -kv_1^{k-1} & -kv_2^{k-1} & -kv_3^{k-1} \\ 1 & 1 & 1 & -1 & -1 & -1 \end{pmatrix}$$

has rank 2, modulo p, and otherwise singular modulo p. Note that for $p \nmid k$, a solution is singular modulo p if and only if

$$(3.34) \qquad u_1^{k-1} \equiv u_2^{k-1} \equiv u_3^{k-1} \equiv v_1^{k-1} \equiv v_2^{k-1} \equiv v_3^{k-1} \bmod p.$$

It follows that there are at most $(k-1)^4 p^{5l-4}$ singular solutions of (3.33), because for each of the p^l choices for u_1, the remaining variables will satisfy (3.34), leaving

at most $(k-1)^4 p^{4l-4}$ possibilities for u_2, u_3, v_1, v_2 by the theory of power residues. The value of $v_3 \bmod p^l$ is then fixed through (3.33).

When $p \nmid k$, we count the non-singular solutions to (3.33) with $u_1^{k-1} \not\equiv v_1^{k-1} \bmod p$ with the aid of Hensel's Lemma. Indeed, each such solution, when reduced modulo p, corresponds to one of the $O(p^4)$ solutions counted by $S_0(p)$, and at most $O(p^{4l-4})$ of them will reduce to the same solution modulo p, because for a solution u, v of the system

$$u^k - v^k \equiv a \bmod p, \quad u - v \equiv b \bmod p$$

with $p \nmid u^{k-1} - v^{k-1}$, there is exactly one pair u_1, v_1 with $1 \le u_1, v_1 \le p^l$ and

$$u_1^k - v_1^k \equiv a \bmod p^l, \quad u_1 - v_1 \equiv b \bmod p^l.$$

A similar argument applies for counting non-singular solutions with $u_1^{k-1} \not\equiv u_2^{k-1} \bmod p$, so that, by symmetry, there are at most $O(p^{4l})$ non-singular solutions of (3.33). It follows that $S_0(p^l) \ll p^{4l} + p^{5l-4}$. We now ignore the condition $p \nmid a$ in (3.29) and use orthogonality to deduce that

$$A(p^l) \ll p^{-4l} S_0(p^l) \ll 1 + p^{l-4}.$$

For $l \ge 2$, this contains (3.31), as required.

We are ready to discuss the first of the two series in (3.28). The easy bound $\mathrm{d}_{j+1}(p^l) \le (l+1)^j$ coupled with (3.30) suffices to recognise the sum

$$(3.35) \qquad \sum_{l=0}^{\infty} \mathrm{d}_{j+1}(p^l) \frac{A(p^l)}{p^l}$$

as a convergent one, and if one uses (3.31) for $p > k \ge l \ge 1$, and (3.30) for $l > k$, then this sum is seen to be of the form $1 + O(p^{-2})$. The sums (3.35) are the factors in the Euler product for the sum (3.28), so that the latter indeed converges.

A similar argument applies to the second sum in (3.28). Rewritten as an Euler product, this sum becomes

$$(3.36) \qquad \prod_{p>Y} \sum_{l=0}^{\infty} j^{L_0 l} \frac{A(p^l)}{p^l}.$$

When P is sufficiently large, one has $j^{L_0} p^{-4/k} < 1/2$ for all $p > Y$, so that (3.30) yields

$$\sum_{l>k} j^{L_0 l} \frac{A(p^l)}{p^l} \ll \sum_{l>k} \left(\frac{j^{L_0}}{p^{4/k}} \right)^l \ll p^{-4}.$$

On using (3.31) for $1 \le l \le k$ as in the previous discussion, one again finds that the Euler factors in (3.36) are of the type $1 + O(p^{-2})$, thus confirming the claims concerning the second sum in (3.28).

4. The principal proposition

Our next result is a version of Hua's lemma with a logarithmic saving, similar to Theorem B of Vaughan [27].

Lemma 8. *Let $k \ge 3$, $K = 2^{k-1}$ and $t = \frac{3}{2}K + 2$. Then*

$$\int_0^1 \int_0^1 |f(\alpha, \beta)|^t \, \mathrm{d}\alpha \, \mathrm{d}\beta \ll P^{t-k-1/2} L^{\varepsilon-3}.$$

Note that Theorem 2 follows on combining the conclusions of Lemmas 1 and 8.

The proof of Lemma 8 will occupy this and the next two sections. Only the initial steps of the proof work for all values of k. Let I denote the integral that is estimated in Lemma 8. On noting that t is the arithmetic mean of $K + 2$ and $2K + 2$, one deduces from Lemma 5 and Schwarz's inequality that

$$(4.1) \qquad I \ll P^{t-k-1/2} L^{\eta}.$$

A similar argument may be applied to the exponential sum $h(\alpha, \beta)$. An application of Schwarz's inequality combined with Lemma 6 yields

$$(4.2) \qquad \int_0^1 \int_0^1 |h(\alpha, \beta)|^t \, d\alpha \, d\beta \ll P^{t-k-1/2} L^{\varepsilon - 1 - \frac{1}{2}(k-1)^2}.$$

We now apply a differencing argument that reduces the estimation of I to that of the integral considered in (4.2). The main ideas are adopted from Vaughan [27].

By orthogonality, I is the number of solutions of the pair of equations

$$(4.3) \qquad x^k - y^k = \sum_{i=1}^{3K/2} (x_i^k - y_i^k), \quad x - y = \sum_{i=1}^{3K/2} (x_i - y_i)$$

with x, y, x_i, y_i all constrained to the interval $[1, P]$. Also, when $\mathscr{U}$ is a subset of $\{(x, y) \in \mathbb{N}^2 : 1 \le x, y \le P\}$, let $I(\mathscr{U})$ denote the number of solutions counted by I that have $(x, y) \in \mathscr{U}$.

Recall the parameters $D = L^\kappa$ and $D' = P^{1/6k}$. For $1 \le x \le P$, let $m(x)$ be the smallest prime factor of x that exceeds D if such a factor exists, and otherwise put $m(x) = \infty$. Consider the sets

$$\mathscr{A} = \{(x, y) \in \mathbb{N}^2 : x \le P, \, y \le P, \, (x; y) > D\},$$

$$\mathscr{B} = \{(x, y) \in \mathbb{N}^2 : x \le P, \, y \le P, \, (x; y) \le D, \, m(x) \le D'\},$$

$$\mathscr{C} = \{(x, y) \in \mathbb{N}^2 : x \le P, \, y \le P, \, (x; y) \le D, \, m(y) \le D'\}.$$

If a solution to (4.3) is counted by I, but the pair (x, y) is not in the union of $\mathscr{A}, \mathscr{B}$ and $\mathscr{C}$, then we have $x \le P$, $y \le P$, $(x; y) \le D$ and $m(x) > D'$, $m(y) > D'$. In particular, $x \in \mathscr{E}$ and $y \in \mathscr{E}$. Consequently

$$I \le I(\mathscr{A}) + I(\mathscr{B}) + I(\mathscr{C}) + I(\mathscr{E} \times \mathscr{E}).$$

Moreover, by symmetry, $I(\mathscr{B}) = I(\mathscr{C})$, and so

$$(4.4) \qquad I \le 4 \max I(\mathscr{X})$$

where $\mathscr{X}$ runs through the sets $\mathscr{A}$, $\mathscr{B}$ and $\mathscr{E} \times \mathscr{E}$.

First suppose that $I \le 4I(\mathscr{E} \times \mathscr{E})$. Then recalling (3.10), orthogonality shows that

$$I \le 4 \int_0^1 \int_0^1 |h(\alpha, \beta)|^2 |f(\alpha, \beta)|^{t-2} \, d\alpha \, d\beta$$

$$\le 4 \left(\int_0^1 \int_0^1 |h(\alpha, \beta)|^t \, d\alpha \, d\beta \right)^{2/t} \left(\int_0^1 \int_0^1 |f(\alpha, \beta)|^t \, d\alpha \, d\beta \right)^{1-2/t}.$$

Here Hölder's inequality was used to infer the second inequality. The second integral on the right is I, and therefore

$$I \le 2^t \int_0^1 \int_0^1 |h(\alpha, \beta)|^t \, d\alpha \, d\beta.$$

Hence in this case, the desired estimate for I is a consequence of (4.2).

Next, suppose that $I \leq 4I(\mathscr{A})$. We define

$$(4.5) \qquad f(\alpha, \beta; W) = \sum_{w \leq W} e(\alpha w^k + \beta w)$$

and sort the pairs $(x, y) \in \mathscr{A}$ according to the value of $d = (x; y)$. Then, by orthogonality, a consideration of the underlying diophantine equations reveals that

$$I(\mathscr{A}) \leq \sum_{d > D} \int_0^1 \int_0^1 |f(\alpha d^k, \beta d; P/d)|^2 |f(\alpha, \beta)|^{t-2} \, d\alpha \, d\beta.$$

The initial assumption and Hölder's inequality yield

$$I \leq 4 \sum_{d > D} \left(\int_0^1 \int_0^1 |f(\alpha d^k, \beta d; P/d)|^t \, d\alpha \, d\beta \right)^{2/t} I^{1-2/t}.$$

We use the bound for I provided by (4.1), and for the remaining integral on the right hand side here, a consideration of the underlying diophantine equations shows that (4.1) again supplies a bound, this time with P/d in place of P. One then finds that

$$I \ll P^{t-k-1/2} L^\eta \sum_{d > D} d^{-(2t-2k-1)/t} \ll P^{t-k-1/2} L^\eta D^{-1/8}.$$

Hence for $\kappa \geq 8(\eta + k^2)$, this shows that $I \ll P^{t-k-1/2} L^{-k^2}$ which is acceptable.

It remains to consider the case where $I \leq 4I(\mathscr{B})$. The initial steps are along familiar lines. Recall the definition of $\mathscr{B}$ and sort the solutions of (4.3) counted by $I(\mathscr{B})$ according to the value of $p = m(x)$. Then $D < p \leq D'$, and the condition that $(x; y) \leq D$ implies $p \nmid y$. Hence $I(\mathscr{B})$ does not exceed the number of solutions to the equation

$$(4.6) \qquad (pw)^k - y^k = \sum_{i=1}^{3K/2} (x_i^k - y_i^k), \quad pw - y = \sum_{i=1}^{3K/2} (x_i - y_i)$$

in primes p with $D < p \leq D'$ and natural numbers w, y, x_i, y_i satisfying $w \leq P/p$ and

$$(4.7) \qquad y \leq P, \quad x_i \leq P, \quad y_i \leq P, \quad p \nmid y.$$

Let I_M denote the number of solutions of (4.6) constrained to (4.7) and $M < p \leq 2M$, $w \leq P/M$. Then, on splitting the range for p into dyadic intervals, one finds that there is some M with $D \leq M \leq D'$ and

$$(4.8) \qquad I \leq 4I(\mathscr{B}) \ll L I_M.$$

We now recall (4.5) and write

$$f_p(\alpha, \beta) = \sum_{\substack{y \leq P \\ p \nmid y}} e(\alpha y^k + \beta y), \qquad g(\alpha, \beta) = f(\alpha, \beta, P/M).$$

Then, by orthogonality,

$$(4.9) \qquad I_M = \sum_{M < p \leq 2M} \int_0^1 \int_0^1 g(p^k \alpha, p\beta) f_p(-\alpha, -\beta) |f(\alpha, \beta)|^{t-2} \, d\alpha \, d\beta.$$

One may reduce the estimation of I to bounding the integral

$$(4.10) \qquad \Theta_p = \int_0^1 \int_0^1 |f_p(\alpha,\beta)|^{K/2} |g(p^k\alpha, p\beta)|^{K+2} \, \mathrm{d}\alpha \, \mathrm{d}\beta,$$

which can be brought into play via Hölder's inequality. Indeed, on noting that $|f_p(-\alpha,-\beta)| = |f_p(\alpha,\beta)|$, one readily finds that

$$\int_0^1 \int_0^1 g(p^k\alpha, p\beta) f_p(-\alpha,-\beta) |f(\alpha,\beta)|^{t-2} \, \mathrm{d}\alpha \, \mathrm{d}\beta$$

$$\leq \Theta_p^{1/(K+2)} I^{1-(2/t)} \Big(\int_0^1 \int_0^1 |f_p(\alpha,\beta)|^t \, \mathrm{d}\alpha \, \mathrm{d}\beta \Big)^{\frac{K+4}{t(2K+4)}}.$$

On considering the underlying diophantine equations, it is immediate that the integral on the far right is bounded by I. Hence

$$\int_0^1 \int_0^1 g(p^k\alpha, p\beta) f_p(-\alpha,-\beta) |f(\alpha,\beta)|^{t-2} \, \mathrm{d}\alpha \, \mathrm{d}\beta \ll \Theta_p^{1/(K+2)} I^{(K+1)/(K+2)}.$$

By orthogonality, the integral on the left is non-negative. We may sum over p to first infer from (4.9) that

$$I_M \ll (IM)^{(K+1)/(K+2)} \Big(\sum_{M < p \leq 2M} \Theta_p \Big)^{1/(K+2)},$$

and then, by (4.8),

$$(4.11) \qquad I \ll L^{K+2} M^{K+1} \sum_{M < p \leq 2M} \Theta_p.$$

The next step is to show that whenever $D \leq M \leq D'$ then

$$(4.12) \qquad \sum_{M < p \leq 2M} \Theta_p \ll P^{t-k-1/2} M^{-K-3/2} L^\eta.$$

Once this is established, it suffices to recall that $D = L^\kappa$, and to combine (4.11) and (4.12) to finally conclude that $I \ll P^{t-k-1/2} L^{-k^2}$ holds in this last case as well, provided only that κ is large enough.

We shall estimate the sum in (4.12) by a differencing argument. When $k \geq 4$, the problem at hand can be approached by combining ideas contained in Vaughan [27] and Wooley [32]. The rather technical details are provided in the next section. For $k = 3$ this argument collapses, and we present an alternative approach via the Hardy-Littlewood method in the following section.

5. Efficient differencing

Throughout this section we suppose that $k \geq 4$. Then $K/4 \geq 2$, and $K/4$ is even. We will use this frequently. We prepare for the differencing operation with a technical estimate concerning certain congruences. For $p > k$, let $\mathscr{Z}_p(a,b)$ be the set of solutions $\mathbf{z} = (z_1, z_2, \ldots, z_{K/4})$ to the simultaneous congruences

$$(5.1) \qquad z_1^k + z_2^k + \ldots + z_{K/4}^k \equiv a \bmod p^k, \quad z_1 + z_2 + \ldots + z_{K/4} \equiv b \bmod p$$

with $1 \leq z_i \leq p^k$ and $p \nmid z_i$ for all $1 \leq i \leq K/4$. Also, let

$$Z_p = \max_{a,b} \# \mathscr{Z}_p(a,b).$$

Lemma 9. *Let $p > k \geq 4$. Then*
$$Z_p \ll p^{\frac{1}{4}Kk - k - 1}.$$

Proof. First suppose that $k = 4$. Then $K/4 = 2$. Consider the solutions w_1, w_2 of

(5.2) $$w_1^4 + w_2^4 \equiv a \bmod p, \quad w_1 + w_2 \equiv b \bmod p$$

with $1 \leq w_i \leq p - 1$ ($i = 1, 2$). Here, one may eliminate w_2. Then w_1 satisfies a polynomial congruence of degree 4 which has at most 4 solutions. For any solution **z** of
$$z_1^4 + z_2^4 \equiv a \bmod p^4, \quad z_1 + z_2 \equiv b \bmod p$$
that is counted by Z_p, there is a solution **w** of (5.2) with $z_i \equiv w_i \bmod p$. However, for a fixed solution **w** of (5.2), there are p^3 choices for z_2 with $1 \leq z_2 \leq p^4$ and $z_2 \equiv w_2 \bmod p$, and for any such z_2, there are at most four solutions of $z_1^4 + z_2^4 \equiv a \bmod p^4$ with $p \nmid z_1$ and $1 \leq z_1 \leq p^4$. This shows that $Z_p \leq 16p^3$, as required.

Next, suppose that $k \geq 5$. Let $\mathscr{Z}_p'(a', b')$ denote the set of solutions z_1, z_2, z_3 of the congruences

(5.3) $$z_1^k + z_2^k + z_3^k \equiv a' \bmod p^k, \quad z_1 + z_2 + z_3 \equiv b' \bmod p$$

with $p \nmid z_1 z_2 z_3$ and $1 \leq z_i \leq p^k$ for $i = 1, 2, 3$. The bound

(5.4) $$\#\mathscr{Z}_p'(a', b') \ll p^{2k-1}$$

holds uniformly for $a, b \in \mathbb{Z}$, $p > k$. This can be seen as follows. In the system of congruences

(5.5) $$w_1^k + w_2^k + w_3^k \equiv a' \bmod p, \quad w_1 + w_2 + w_3 \equiv b' \bmod p$$

one may eliminate w_3. But $w_1^k + w_2^k + (b' - w_1 - w_2)^k$ is a polynomial in w_1, and when $2 \mid k$, the degree is k and the leading coefficient is 2. Hence we find at most kp incongruent solutions of (5.5). When k is odd, and $w_2 \not\equiv b' \bmod p$, then the degree is $k - 1$ and the leading coefficient is $k(b' - w_2)$, so that (5.5) can have at most $(k-1)(p-1)$ solutions with $w_2 \not\equiv b' \bmod p$, and further p solutions with $w_2 \equiv b' \bmod p$. Hence in all cases, there are at most kp solutions. Any solution of (5.3) reduces to one of (5.5). There are p^{2k-2} choices of $z_2, z_3 \bmod p^k$ with $z_i \equiv w_i \bmod p$. Now solve for z_1 from the first congruence in (5.3). Since $p \nmid z_1$, there are at most k solutions for z_1. This confirms (5.4). To complete the proof of the lemma, it now suffices to take
$$a' = a - z_4^k - \ldots - z_{K/4}^k, \quad b' = b - z_4 - \ldots - z_{K/4}$$
in (5.4), and to sum over $z_4, \ldots, z_{K/4}$ trivially.

We now return to the main theme. Let p be a prime with $M < p \leq 2M$. The goal is to estimate Θ_p, as defined in (4.10). By orthogonality, Θ_p is the number of solutions to the pair of equations

(5.6) $$\sum_{i=1}^{K/4}(x_i^k - y_i^k) = p^k \sum_{i=1}^{1+K/2}(u_i^k - v_i^k), \quad \sum_{i=1}^{K/4}(x_i - y_i) = p \sum_{i=1}^{1+K/2}(u_i - v_i)$$

in natural numbers x_i, y_i, u_i, v_i constrained to

(5.7) $$x_i \leq P, \quad y_i \leq P, \quad u_i \leq P/M, \quad v_i \leq P/M, \quad p \nmid x_i y_i \quad (1 \leq i \leq K/4).$$

Note that solutions to (5.6) satisfy

(5.8)
$$\begin{aligned}
x_1^k + x_2^k + \ldots + x_{K/4}^k &\equiv y_1^k + y_2^k + \ldots + y_{K/4}^k \bmod p^k, \\
x_1 + x_2 + \ldots + x_{K/4} &\equiv y_1 + y_2 + \ldots + y_{K/2} \bmod p.
\end{aligned}$$

The following argument is an elementary variant of a very similar exponential sum technique underpinning the proof of Lemma 2.2 in Wooley [32]. For given data $n, m, z_1, \ldots, z_{K/4}$, let $\Phi_p(\mathbf{z}, n, m)$ denote the number of solutions to

$$\sum_{i=1}^{K/4} x_i^k + p^k \sum_{i=1}^{1+K/2} u_i^k = n, \qquad \sum_{i=1}^{K/4} x_i + p \sum_{i=1}^{1+K/2} u_i = m$$

satisfying the relevant conditions in (5.7) and $x_i \equiv z_i \bmod p^k$, for $1 \leq i \leq K/4$. Then, by (5.8) and the discussion preceding this observation,

$$\Theta_p = \sum_{n,m=-\infty}^{\infty} \sum_{a=1}^{p^k} \sum_{b=1}^{p} \Big| \sum_{\mathbf{z} \in \mathscr{Z}_p(a,b)} \Phi_p(\mathbf{z}, n, m) \Big|^2.$$

By Cauchy's inequality and Lemma 9,

$$\Theta_p \leq \sum_{n,m=-\infty}^{\infty} \sum_{a=1}^{p^k} \sum_{b=1}^{p} Z_p \sum_{\mathbf{z} \in \mathscr{Z}_p(a,b)} \big| \Phi_p(\mathbf{z}, n, m) \big|^2$$

(5.9)
$$\ll p^{\frac{1}{4} Kk - k - 1} \sum_{n,m=-\infty}^{\infty} \sum_{\mathbf{z} \bmod p^k} \big| \Phi_p(\mathbf{z}, n, m) \big|^2 \ll p^{\frac{1}{4} Kk - k - 1} \Psi_p$$

where Ψ_p is the number of solutions to (5.6) and (5.7) with the additional constraints that

(5.10)
$$x_i \equiv y_i \bmod p^k \quad (1 \leq i \leq K/4).$$

Now let

(5.11)
$$G(\alpha, \beta) = \sum_{\substack{x \leq P, y \leq P \\ x \equiv y \bmod p^k}} e\big(\alpha(x^k - y^k) + \beta(x - y)\big).$$

By orthogonality,

(5.12)
$$\Psi_p \leq \int_0^1 \int_0^1 G(\alpha, \beta)^{K/4} |g(\alpha p^k, \beta p)|^{K+2} \, d\alpha \, d\beta.$$

By (5.11),

$$G(\alpha, \beta) = [P] + 2\mathrm{Re} \sum_{\substack{y < x \leq P \\ x \equiv y \bmod p^k}} e\big(\alpha(x^k - y^k) + \beta(x - y)\big)$$

which implies that

(5.13)
$$|G(\alpha, \beta)|^{K/4} \ll P^{K/4} + \Big| \sum_{\substack{y < x \leq P \\ x \equiv y \bmod p^k}} e\big(\alpha(x^k - y^k) + \beta(x - y)\big) \Big|^{K/4}.$$

We use this in (5.12) and apply Lemma 5 to deduce that

(5.14)
$$\Psi_p \ll \Psi_p' + P^{K/4} (P/M)^{K+2-k} L^\eta$$

where
$$\Psi'_p = \int_0^1 \int_0^1 \Big| \sum_{\substack{y < x \le P \\ x \equiv y \bmod p^k}} e\big(\alpha(x^k - y^k) + \beta(x - y)\big)\Big|^{K/4} |g(p^k\alpha, p\beta)|^{K+2} \, d\alpha \, d\beta.$$

By orthogonality, Ψ'_p is the number of solutions of the pair of equations
$$\sum_{i=1}^{K/4}(-1)^i(x_i^k - y_i^k) = p^k \sum_{i=1}^{1+K/2}(u_i^k - v_i^k), \quad \sum_{i=1}^{K/4}(-1)^i(x_i - y_i) = p \sum_{i=1}^{1+K/2}(u_i - v_i)$$

subject to (5.7), (5.10) and $x_i > y_i$. We write
$$z_i = x_i + y_i, \quad h_i = p^{-k}(x_i - y_i).$$

Then, since $2x_i = z_i + p^k h_i$ and $2y_i = z_i - p^k h_i$, one finds that
$$2^k(x_i^k - y_i^k) = p^k \varphi_p(z_i, h_i)$$

where φ_p is the integral polynomial
$$\varphi_p(z, h) = p^{-k}\big((z + hp^k)^k - (z - hp^k)^k\big),$$

and we may then conclude that Ψ'_p does not exceed the number of solutions of the system
$$\sum_{i=1}^{K/4}(-1)^i\varphi_p(z_i, h_i) = 2^k \sum_{i=1}^{1+K/2}(u_i^k - v_i^k), \qquad p^{k-1}\sum_{i=1}^{K/4}(-1)^i h_i = \sum_{i=1}^{1+K/2}(u_i - v_i)$$

in which the variables are subject to
$$1 \le h_i \le PM^{-k}, \quad z_i \le 2P, \quad u_i \le P/M, \quad v_i \le P/M.$$

Now write $H = PM^{-k}$ and introduce the exponential sum
$$F_p(\alpha, \beta) = \sum_{h \le H} \sum_{z \le 2P} e\big(\alpha\varphi_p(z, h) + \beta h p^{k-1}\big).$$

Then, once again by orthogonality,
$$(5.15) \qquad \Psi'_p \le \int_0^1 \int_0^1 |F_p(\alpha, \beta)|^{K/4} |g(2^k\alpha, \beta)|^{K+2} \, d\alpha \, d\beta.$$

We have now completed the first differencing step. The differencing was efficient because the congruences (5.8) reduce the potential reservoir for the variables x_i, y_i by a factor p^{-k-1}, and one recovers this through Lemma 9. We proceed by taking further differences. By Cauchy's inequality,
$$|F_p(\alpha, \beta)|^2 \le H \sum_{h \le H} \Big| \sum_{z \le 2P} e\big(\alpha\varphi_p(z, h)\big)\Big|^2.$$

Note that β is absent from the right hand side. Thus, we may proceed as with the usual proof of Weyl's inequality (see Lemma 2.4 of Vaughan [29]) to confirm that there are certain natural numbers $c(l)$ (depending also on p which we suppress) with
$$c(0) \ll P^{k-1}M^{-k}, \qquad c(l) \ll d_k(l) \quad (l \ne 0),$$

and such that
$$(5.16) \qquad |F_p(\alpha, \beta)|^{K/2} \le P^{K-k} M^{(1-\frac{1}{2}K)k} \sum_l c(l)e(\alpha l).$$

At this point it might be worth recalling that in the current context one has $k \geq 4$, and that therefore β is indeed absent from this bound, for all k under consideration. To bound the integral

$$V = \int_0^1 \int_0^1 |F_p(\alpha, \beta)|^{K/2} |g(2^k \alpha, \beta)|^{K+2} \, \mathrm{d}\alpha \, \mathrm{d}\beta,$$

one inserts the inequality (5.16), and then separates off the contribution arising from $l = 0$. This term will contribute to V at most

$$\ll P^{K-k} M^{(1-\frac{1}{2}K)k} c(0) \int_0^1 \int_0^1 |g(2^k \alpha, \beta)|^{K+2} \, \mathrm{d}\alpha \, \mathrm{d}\beta.$$

We may apply Lemma 5 to the integral on the right, and then conclude that this contribution to V does not exceed

$$\ll (P^{K-k} M^{(1-\frac{1}{2}K)k})(P^{k-1} M^{-k})(P/M)^{K+2-k} L^\eta \ll P^{2K-k+1} M^{(1-\frac{1}{2}K)k-K-2} L^\eta.$$

By orthogonality, the terms that correspond to $l \neq 0$ produce a term not exceeding

$$\ll P^{K-k} M^{(1-\frac{1}{2}K)k} \sum_{u_i, v_i} \mathrm{d}_k\big(u_1^k - v_1^k + \ldots + u_{\frac{1}{2}K+1}^k - v_{\frac{1}{2}K+1}^k\big)$$

in which the variables are restricted by $u_i \leq P/M$, $v_i \leq P/M$ and

$$u_1 - v_1 + \ldots + u_{\frac{1}{2}K+1} - v_{\frac{1}{2}K+1} = 0.$$

By Lemma 3, this does not exceed

$$\ll (P^{K-k} M^{(1-\frac{1}{2}K)k})(P/M)^{K+1} L^\eta \ll P^{2K-k+1} M^{(1-\frac{1}{2}K)k-1-K} L^\eta.$$

Collecting together yields

$$V \ll P^{2K-k+1} M^{(1-\frac{1}{2}K)k-1-K} L^\eta.$$

By (5.15) and Hölder's inequality,

$$\Psi_p' \leq V^{1/2} \Big(\int_0^1 \int_0^1 |g(2^k \alpha, \beta)|^{K+2} \, \mathrm{d}\alpha \, \mathrm{d}\beta \Big)^{1/2},$$

and another use of Lemma 5 produces

(5.17) $$\Psi_p' \ll P^{\frac{3}{2}K-k+\frac{3}{2}} M^{\frac{1}{2}(1-\frac{1}{2}K)k-K-\frac{3}{2}+\frac{1}{2}k} L^\eta.$$

By (5.9) and (5.14), we may now conclude that

$$\sum_{M < p \leq 2M} \Theta_p \ll M^{\frac{1}{4}Kk-k} (\max_p \Psi_p' + P^{K/4}(P/M)^{K+2-k} L^\eta).$$

As is readily checked, this establishes (4.12), as was required to complete the proof of Lemma 8.

6. Inefficient differencing

In this section we establish the case $k = 3$ of Lemma 8. Our approach needs substantial revision because Lemma 4 provides optimal control on the sixth moment of Weyl sums, and if one differences two blocks of two variables beyond this as would be needed for efficient differencing, then one works with ten variables. But when $k = 3$ one has $t = 8$. One could, at least in principle, study a tenth moment, but savings can then be expected only if differencing is performed over minor arcs only. This would entail considerable complication in detail, and we prefer an eighth moment for consistency with our work in the previous section. Fortunately, the

inevitable loss of a factor M in an inefficient differencing can be restored in part by averaging over the auxiliary prime p. The technique elaborates on ideas of Vaughan [26].

We now return to (4.10), temporarily fix a prime p with $M < p \le 2M$, and observe by orthogonality that Θ_p equals the number of solutions of the diophantine system

$$(6.1) \qquad \begin{aligned} x_1^3 + p^3(y_1^3 + y_2^3 + y_3^3) &= x_2^3 + p^3(y_4^3 + y_5^3 + y_6^3), \\ x_1 + p(y_1 + y_2 + y_3) &= x_2 + p(y_4 + y_5 + y_6) \end{aligned}$$

with variables constrained to

$$x_i \le P, \quad y_i \le P/M, \quad p \nmid x_1 x_2.$$

Any solution of (6.1) satisfies $x_1^3 \equiv x_2^3 \bmod p^3$ and $x_1 \equiv x_2 \bmod p$. Since $p \nmid x_1 x_2$, this implies $x_1 \equiv x_2 \bmod p^3$. By Lemma 4, the number of solutions with $x_1 = x_2$ amount to at most $O(P(P/M)^3)$. By symmetry, it now suffices to count solutions where $x_1 > x_2$. According to the preceding comment, we put $x_1 = x_2 + hp^3$ with $h > 0$, and $z = x_1 + x_2$. Then (6.1) transforms to

$$(6.2) \qquad \begin{aligned} h(3z^2 + h^2 p^6) &= 4(y_1^3 + y_2^3 + y_3^3 - y_4^3 - y_5^3 - y_6^3), \\ hp^2 &= y_1 + y_2 + y_3 - y_4 - y_5 - y_6. \end{aligned}$$

Let Ξ denote the number of solutions of (6.2) subject to

$$z \le 2P, \quad y_i \le P/M, \quad h \le H, \quad M < p \le 2M$$

where $H = PM^{-3}$. Then, on summing over p, the above argument yields

$$(6.3) \qquad \sum_{M < p \le 2M} \Theta_p \le 2\Xi + O(P^4 M^{-2}).$$

Let

$$F(\alpha) = \sum_{z \le 2P} e(3\alpha z^2), \quad E(\alpha, \beta) = \sum_{M < p \le 2M} e(\alpha p^6 + \beta p^2)$$

and recall that $g(\alpha, \beta) = f(\alpha, \beta, P/M)$. Then, by orthogonality,

$$(6.4) \qquad \Xi = \int_0^1 \int_0^1 \sum_{h \le H} F(\alpha h) E(\alpha h^3, \beta h) |g(4\alpha, \beta)|^6 \, d\alpha \, d\beta.$$

Note the similarity with (5.15).

We apply the Hardy-Littlewood method to estimate Ξ. Let $\delta > 0$ be a small parameter to be determined later. Let $\mathfrak{N}$ be the set of all $\alpha \in [0, 1]$ where there are coprime a, q with $0 \le a \le q$, $1 \le q \le P^{1+\delta}$ and $|q\alpha - a| \le P^{\delta - 1} H^{-1}$. Let $\mathfrak{n}$ be the complement of $\mathfrak{N}$ in $[0, 1]$. Then, by Lemma 4 of Vaughan [26], one has

$$(6.5) \qquad \sup_{\alpha \in \mathfrak{n}} \sum_{h \le H} |F(\alpha h)|^2 \ll HPL^{3/2}.$$

Unlike in [26], the actual exponent of L is of no importance for us. Next recall (4.5) to observe that

$$\sum_{h \le H} |E(\alpha h^3, \beta h)|^2 = \sum_{M < p_1, p_2 \le 2M} f\big(\alpha(p_1^6 - p_2^6), \beta(p_1 - p_2), H\big).$$

The contribution of summands with $p_1 = p_2$ here is $O(HM)$. We claim that whenever $\alpha \in \mathfrak{n}$ and $p_1 \neq p_2$, then $f\big(\alpha(p_1^6 - p_2^6), \beta(p_1 - p_2), H\big) \ll H^{3/4}(\log H)^2$. Once this is established, it follows that

$$(6.6) \qquad \sum_{h \leq H} |E(\alpha h^3, \beta h)|^2 \ll HM + M^2 H^{3/4}(\log H)^2 \ll HM$$

holds for $M \leq P^{1/15}$. To confirm these claims, let $M < p_1, p_2 \leq 2M$ with $p_1 \neq p_2$ and $|f\big(\alpha(p_1^6 - p_2^6), \beta(p_1 - p_2), H\big)| > H^{3/4}(\log H)^2$. Then, by Lemma 1 (with $\theta = 7/9$), there are coprime numbers b, r with $r \leq H^{7/9}$ and $|\alpha r(p_1^6 - p_2^6) - b| \leq H^{-20/9}$, and consequently there is a q with $q \mid r(p_1^6 - p_2^6)$ and an $a \in \mathbb{Z}$ with $|\alpha q - a| \leq H^{-20/9}$. Hence $q \leq H^{7/9}(2M)^6 = 2^6 P^{7/9} M^{11/3}$, and therefore $\alpha \in \mathfrak{N}$, as one readily confirms, and as was desired. This establishes (6.6).

By (6.5), (6.6) and Cauchy's inequality,

$$\sup_{\alpha \in \mathfrak{n}} \Big| \sum_{h \leq H} F(\alpha h) E(\alpha h^3, \beta h) \Big| \ll H(PM)^{1/2} L.$$

We use this in conjunction with Lemma 4 to infer that

$$(6.7) \qquad \int_{\mathfrak{n}} \int_0^1 \sum_{h \leq H} F(\alpha h) E(\alpha h^3, \beta h) |g(4\alpha, \beta)|^6 \, \mathrm{d}\beta \, \mathrm{d}\alpha \ll H(PM)^{1/2}(P/M)^3 L.$$

We now consider the major arcs $\mathfrak{N}$. When a, q are coprime with $|q\alpha - a| \leq 1/q$, then Lemma 3.1 of Vaughan [28] asserts that

$$\sum_{h \leq H} |F(\alpha h)|^2 \ll P^\varepsilon \Big(\frac{P^2 H}{q + P^2 H |q\alpha - a|} + PH + q + P^2 H |q\alpha - a| \Big).$$

For $\alpha \in \mathfrak{N}$, there are such a, q with $q \leq P^{1+\delta}$ and $|q\alpha - a| \leq H^{-1} P^{\delta - 1}$, and a short calculation then confirms that

$$(6.8) \qquad \sum_{h \leq H} |F(\alpha h)|^2 \ll \frac{P^{2+3\delta} H}{q + P^2 H |q\alpha - a|}.$$

Moreover, by orthogonality,

$$(6.9) \qquad \int_0^1 |g(4\alpha, \beta)|^4 \, \mathrm{d}\beta = \sum_l \psi_l e(\alpha l)$$

in which ψ_l equals the number of solutions of the pair of diophantine equations

$$4(y_1^3 + y_2^3 - y_3^3 - y_4^3) = l, \quad y_1 + y_2 - y_3 - y_4 = 0$$

with $y_i \leq P/M$ $(1 \leq i \leq 4)$. The integral representation shows that the Fourier series in (6.9) takes non-negative values only. Hence we may apply Lemma 2 of Brüdern [5] to conclude that

$$\int_{\mathfrak{N}} \sum_{h \leq H} |F(\alpha h)|^2 \int_0^1 |g(4\alpha, \beta)|^4 \, \mathrm{d}\beta \, \mathrm{d}\alpha \ll P^{3\delta + \varepsilon} \Big(P^{1+\delta} \psi_0 + \sum_l \psi_l \Big),$$

and the bounds $\psi_0 \ll (P/M)^2$ and $\sum_l \psi_l \ll (P/M)^3$ are immediate. This gives

$$(6.10) \qquad \int_{\mathfrak{N}} \sum_{h \leq H} |F(\alpha h)|^2 \int_0^1 |g(4\alpha, \beta)|^4 \, \mathrm{d}\beta \, \mathrm{d}\alpha \ll P^{3+5\delta} M^{-2}.$$

Also, on using a trivial bound for $E(\alpha, \beta)$ and Lemma 5 only, we have

$$\int_{\mathfrak{N}} \int_0^1 \sum_{h \le H} |E(\alpha h^3, \beta h)|^2 |g(4\alpha, \beta)|^8 \, d\beta \, d\alpha$$

$$(6.11) \qquad \ll HM^2 \left(\int_0^1 \int_0^1 |g(4\alpha, \beta)|^{10} \, d\alpha \, d\beta \right)^{1/2} \left(\int_0^1 \int_0^1 |g(4\alpha, \beta)|^6 \, d\alpha \, d\beta \right)^{1/2}$$

$$\ll HM^2 (P/M)^{9/2} L^\eta.$$

By (6.10), (6.11) and Schwarz's inequality,

$$\int_{\mathfrak{N}} \int_0^1 \sum_{h \le H} F(\alpha h) E(\alpha h^3, \beta h) |g(4\alpha, \beta)|^6 \, d\beta \, d\alpha \ll P^{17/4 + 3\delta} M^{-15/4}.$$

When $M \le P^{1/15}$, this bound is superior to the one in (6.7). Hence, by (6.7) and (6.4) we conclude that $\Xi \ll P^{9/2} M^{-11/2} L$. The case $k = 3$ of Lemma 8 now follows via (6.3).

7. Pruning

In order to facilitate the major arc analysis within the proof of Theorem 1, the minor arc estimate provided by Theorem 2 is to be augmented by a device that restricts the integration for the linear equation to suitable major arcs as well. Some notation is required to make this precise.

Let $1 \le q \le P$, and let $\mathfrak{M}_q$ denote the union of the intervals

$$\{\alpha \in [0, 1] : |q\alpha - a| \le (4k)^{-1} P^{1-k}\}$$

with $0 \le a \le q$ and $(a; q) = 1$. Note that these intervals are disjoint, as are the various $\mathfrak{M}_q$. Let $\mathfrak{K}_q = \mathfrak{M}_q \times [0, 1]$, and let $\mathfrak{K}$ be the disjoint union of the $\mathfrak{K}_q$ with $1 \le q \le P$. Further, when $1 \le q \le P^{1/9}$, let $\mathfrak{L}_q$ be the union of the rectangles

$$\left\{ (\alpha, \beta) \in [0, 1]^2 : \left| \alpha - \tfrac{a}{q} \right| \le P^{1/9-k}, \left| \beta - \tfrac{b}{q} \right| \le P^{-8/9} \right\}$$

with $0 \le a \le q$, $0 \le b \le q$ and $(a; q) = 1$. Then, for $1 \le q \le P^{1/9}$, one has $\mathfrak{L}_q \subset \mathfrak{K}_q$. Let $\mathfrak{L}$ be the union of $\mathfrak{L}_q$ with $1 \le q \le P^{1/9}$ so that $\mathfrak{L} \subset \mathfrak{K}$.

Lemma 10. *Let $s > 3k$. Then*

$$\iint_{\mathfrak{K} \backslash \mathfrak{L}} |f(\alpha, \beta)|^s \, d\alpha \, d\beta \ll P^{s-k-1-1/(9k)+\varepsilon}.$$

Proof. Let $(\alpha, \beta) \in \mathfrak{K}$. Then there are a unique $q \in [1, P]$ with $(\alpha, \beta) \in \mathfrak{K}_q$, and unique a, b with $0 \le a \le q$, $0 \le b \le q$ with $(a; q) = 1$ and

$$|q\alpha - a| \le (4k)^{-1} P^{1-k}, \quad -\tfrac{1}{2} < q\beta - b \le \tfrac{1}{2}.$$

Define $f^* : \mathfrak{K} \to \mathbb{C}$ by

$$f^*(\alpha, \beta) = q^{-1} S(q, a, b) v\left(\alpha - \tfrac{a}{q}, \beta - \tfrac{b}{q} \right).$$

Whenever $(\alpha, \beta) \in \mathfrak{K}$, one infers from Theorem 3 that

$$f(\alpha, \beta) = f^*(\alpha, \beta) + O\left(P^{(k-1)/k+\varepsilon} \right).$$

Consequently

$$|f(\alpha, \beta)|^s \ll |f^*(\alpha, \beta)|^s + P^{s-s/k+\varepsilon}.$$

The linear measure of $\mathfrak{M}_q$ is $O(P^{1-k})$, whence the planar measure of $\mathfrak{K}$ is $O(P^{2-k})$. Since $s \geq 3k + 1$, it follows that

$$(7.1) \qquad \iint_{\mathfrak{K} \setminus \mathfrak{L}} |f(\alpha, \beta)|^s \, \mathrm{d}\alpha \, \mathrm{d}\beta \ll \iint_{\mathfrak{K} \setminus \mathfrak{L}} |f^*(\alpha, \beta)|^s \, \mathrm{d}\alpha \, \mathrm{d}\beta + P^{s-k-1-1/k+\varepsilon}.$$

Now consider the integral on the right hand side. First, we estimate the contribution from $\mathfrak{K}_q$ with $P^{1/9} < q \leq P$ which amounts to

$$(7.2) \qquad \sum_{\substack{P^{1/9} < q \leq P}} \sum_{\substack{a=1 \\ (a;q)=1}}^{q} \sum_{b=1}^{q} q^{-s} |S(q, a, b)|^s \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |v(\xi, \zeta)|^s \, \mathrm{d}\xi \, \mathrm{d}\zeta.$$

By (2.2), we have the alternative estimates

$$(7.3) \qquad v(\xi, \zeta) \ll P(1 + P^k |\xi|)^{-1/k}, \quad v(\xi, \zeta) \ll P(1 + P|\zeta|)^{-1/k},$$

and consequently

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |v(\xi, \zeta)|^s \, \mathrm{d}\xi \, \mathrm{d}\zeta \ll P^{s-k-1}.$$

By (2.2) again, the expression in (7.2) is bounded by

$$\ll \Big(\sum_{P^{1/9} < q \leq P} q^{2-s/k+\varepsilon} \Big) P^{s-k-1} \ll P^{s-k-1-(1/9k)+\varepsilon}$$

which is acceptable.

It remains to consider the contribution from the sets $\mathfrak{K}_q \setminus \mathfrak{L}_q$ with $q \leq P^{1/9}$ to the integral on the right hand side of (7.1), which in fact does not exceed

$$(7.4) \qquad \sum_{q \leq P^{1/9}} \sum_{\substack{a=1 \\ (a;q)=1}}^{q} \sum_{b=1}^{q} q^{-s} |S(q, a, b)|^s (B_1 + B_2)$$

where

$$B_1 = \int_{-\infty}^{\infty} \int_{|\xi| > P^{1/9-k}} |v(\xi, \zeta)|^s \, \mathrm{d}\xi \, \mathrm{d}\zeta, \quad B_2 = \int_{-\infty}^{\infty} \int_{|\zeta| > P^{-8/9}} |v(\xi, \zeta)|^s \, \mathrm{d}\zeta \, \mathrm{d}\xi.$$

By (7.3), and since $s \geq 3k + 1$, one has

$$(7.5)$$
$$B_1 \ll P^s \int_{-\infty}^{\infty} (1 + P|\zeta|)^{-1-1/k} \, \mathrm{d}\zeta \int_{|\xi| > P^{1/9-k}} (2 + P^k |\xi|)^{-2} \, \mathrm{d}\xi \ll P^{s-k-10/9}.$$

A similar reasoning yields the same bound for B_2. By (2.2), it is now readily seen that the expression in (7.4) does not exceed $O(P^{s-k-10/9})$, and the lemma follows from (7.1).

8. The proof of Theorem 1

This final section is devoted to the proof of Theorem 1 which we launch by disposing of some simple cases. Throughout this section, let $s \geq 2^k + 2$, let $a_i \neq 0$ for $1 \leq i \leq s$, and suppose that exactly r of the s numbers b_i are non-zero. Since the equations (1.1) have a non-singular real solution, one concludes that $r \geq 1$. On

renumbering the variables, we may arrange that $b_i \neq 0$ for $1 \leq i \leq r$. The system (1.1) now takes the shape

$$
\begin{array}{ll}
(8.1) & \begin{aligned}
a_1 x_1^k + \cdots + a_r x_r^k \quad + \cdots + a_s x_s^k &= 0, \\
b_1 x_1 + \cdots + b_r x_r \quad\quad\quad\quad\quad &= 0.
\end{aligned}
\end{array}
$$

Here, we may clear common factors in the second equation, so that we may further suppose that $\mathbf{b}$ is a primitive vector.

If $r = 1$ then $x_1 = 0$, and we are left with a single equation in $2^k + 1$ variables that may be treated by the methods of chapter 2 of Vaughan [29].

If $r = 2$ and $(b_1; b_2) = 1$, then any solutions to (8.1) has $b_1 \mid x_2$, $b_2 \mid x_1$. We substitute $x_1 = b_2 y$, $x_2 = -b_1 y$ to infer that $N(P)$ equals the number of solutions to the equation

$$
(a_1 b_2^k - a_2 b_1^k) y^k + a_3 x_3^k + \ldots a_s x_s^k = 0
$$

with $|x_i| \leq P$ and $|y| \leq P / \max(|b_1|, |b_2|)$. If $a_1 b_2^k \neq a_2 b_1^k$, then this single equation in $s - 1 \geq 2^k + 1$ variables may again be treated by the methods of chapter 2 of Vaughan [29], the unconventional size constrained on y being readily absorbed by the classical method that need not be commented on any further here. If $a_1 b_2^k = a_2 b_1^k$, one has

$$
N(P) = \big(2P / \max(|b_1|, |b_2|) + O(1)\big) N_0
$$

where N_0 is the number of solutions to the equation $a_3 x_3^k + \cdots + a_s x_s^k = 0$ with $|x_i| \leq B$. Now $s - 2 \geq 2^k$. If $s - 2 > 2^k$, then we may proceed as before to establish an asymptotic formula for N_0 by the methods of chapter 2 of Vaughan [29], and this will complete the proof of Theorem 1 in this case. This leaves the case where $s - 2 = 2^k$. In this case, the methods of Vaughan [27] (and in particular (1.5)) may be combined with the singular series work in chapter 2 and 4 of Vaughan [29] to establish that

$$
N_0 = \chi_\infty \Big(\prod_p \chi_p \Big) P^{s-k-2} \big(1 + O((\log P)^{-1})\big)
$$

where the Euler product is absolutely convergent. Also, χ_p is non-zero if and only if $a_3 x_3^k + \cdots + a_s x_s^k = 0$ has a no trivial solution in $\mathbb{Q}_p$, and χ_∞ is non-zero if and only if $a_3, a_4, \ldots, a_s$ are not all of the same sign. On collecting together, this establishes Theorem 1 in the case $r = 2$.

This leaves the case $r \geq 3$. Here we define

$$
(8.2) \qquad \mathscr{F}(\alpha, \beta) = \prod_{i=1}^{s} f(a_i \alpha, b_i \beta),
$$

and observe that the integral

$$
(8.3) \qquad N_{\mathbf{a}, \mathbf{b}}^{+}(P) = \int_0^1 \int_0^1 \mathscr{F}(\alpha, \beta) \, d\alpha \, d\beta
$$

counts the solutions of (1.1) with $1 \leq x_i \leq P$ ($i = 1, \ldots, s$).

In preparation for an application of the Hardy-Littlewood method to the integral (8.3), we define the major arcs $\mathfrak{W}$ as the union of the boxes

$$
(8.4) \qquad \{(\alpha, \beta) \in [0,1]^2 \colon \big|\alpha - \tfrac{a}{q}\big| \leq P^{(1/8)-k}, \ \big|\beta - \tfrac{b}{q}\big| \leq P^{-7/8}\}
$$

with $0 \leq a \leq q$, $0 \leq b \leq q$, $(a; b; q) = 1$ and $q \leq P^{1/8}$. This union is disjoint. When $(\alpha, \beta) \in \mathfrak{W}$ is in the box (8.4), we put

$$
f_i^*(\alpha, \beta) = q^{-1} S(q, a a_i, b b_i) v\big(a_i(\alpha - a/q), b_i(\beta - b/q)\big).
$$

Note that Theorem 3 is not a suitable tool to compare $f(a_i\alpha, b_i\beta)$ with $f_i^*(\alpha, \beta)$ because the condition $(a; q) = 1$ is not met on some boxes. However, Theorem 7.2 of Vaughan [29] readily yields

$$f(a_i\alpha, b_i\beta) = f_i^*(\alpha, \beta) + O(P^{1/4+\varepsilon})$$

uniformly for $(\alpha, \beta) \in \mathfrak{W}$. When inserted into (8.2), the "trivial" bounds $f(\alpha, \beta) \ll P$ (see (1.4)) and $f_i^*(\alpha, \beta) \ll P$ (see (2.2)) suffice to infer that

$$(8.5) \qquad \mathscr{F}(\alpha, \beta) = \prod_{i=1}^{s} f_i^*(\alpha, \beta) + O(P^{s-3/4+\varepsilon}).$$

We put

$$(8.6) \qquad T(q) = \sum_{\substack{a,b=1 \\ (a;b;q)=1}}^{q} q^{-s} \prod_{i=1}^{s} S(q, aa_i, bb_i)$$

and integrate (8.5) over $\mathfrak{W}$. Since the measure of $\mathfrak{W}$ is $O(P^{-k-3/8})$, this yields

$$(8.7) \qquad \iint_{\mathfrak{W}} \mathscr{F}(\alpha, \beta) \, \mathrm{d}\alpha \, \mathrm{d}\beta = \sum_{q \leq P^{1/8}} T(q) \int_{-P^{1/8-k}}^{P^{1/8-k}} \int_{-P^{-7/8}}^{P^{-7/8}} \prod_{i=1}^{s} v(a_i\xi, b_i\zeta) \, \mathrm{d}\zeta \, \mathrm{d}\xi$$
$$+ O(P^{s-k-9/8+\varepsilon}).$$

Note that the main term on the right hand side of (8.7) is a product. The next natural step is to complete the sum over q to a series, and likewise, to complete the integral to one extended over $\mathbb{R}^2$. This requires some care in cases where many of the b_i are zero. In fact, when $r = 3$, the bounds in (2.2) are not of strength sufficient to conclude that the integrand in (8.7) is integrable over $\mathbb{R}^2$. In preparation for a debugging argument, let

$$v_0(\xi, \zeta) = \int_0^1 e(\xi t^k + \zeta t) \, \mathrm{d}t.$$

By (2.1), one has $v(\xi, \zeta) = Pv_0(P^k\xi, P\zeta)$ so that an obvious substitution gives

$$(8.8) \qquad \int_{-P^{1/8-k}}^{P^{1/8-k}} \int_{-P^{-7/8}}^{P^{-7/8}} \prod_{i=1}^{s} v(a_i\xi, b_i\zeta) \, \mathrm{d}\zeta \, \mathrm{d}\xi = P^{s-k-1} \mathscr{J}(P^{1/8})$$

where

$$(8.9) \qquad \mathscr{J}(W) = \int_{-W}^{W} \int_{-W}^{W} \prod_{i=1}^{s} v_0(a_i\xi, b_i\zeta) \, \mathrm{d}\zeta \, \mathrm{d}\xi.$$

If $\zeta \geq 1$ and $\zeta \geq 8k|\xi|$, then $\frac{\mathrm{d}}{\mathrm{d}t}(\xi t^k + \zeta t) \geq \frac{3}{4}\zeta$ holds for all $0 \leq t \leq 1$. Observing symmetry, Lemma 4.2 of Titchmarsh [25] now shows that

$$(8.10) \qquad v_0(\xi, \zeta) \ll (1 + |\zeta|)^{-1}$$

holds throughout the domain described by $|\zeta| \geq 8k|\xi|$. It follows that

$$(8.11) \qquad \prod_{i=1}^{s} v_0(a_i\xi, b_i\zeta) \ll \begin{cases} (1 + |\zeta|)^{-2}(1 + |\xi|)^{-2} & \text{if } |\zeta| \geq 8k|\xi|, \\ (1 + |\xi|)^{-3} & \text{if } |\zeta| < 8k|\xi|. \end{cases}$$

To see this, first note that (7.3) (with $P = 1$) yields $v_0(a_i\xi, b_i\zeta) \ll (1 + |\xi|)^{-1/k}$ because $|a_i| \geq 1$. Since $s > 3k$, this already shows the left hand side of (8.11)

bounded by $(1 + |\xi|)^{-3}$ irrespective of the value for ζ. If $|\zeta| \geq 8k|\xi|$, we estimate $v_0(a_i\xi, b_i\zeta)$ as before for $3 \leq i \leq s$, and this yields a factor $(1 + |\xi|)^{-2}$ for the upper bound. Now recall that $b_1 b_2 \neq 0$, so that (8.10) gives $v_0(a_1\xi, b_1\zeta)v_0(a_2\xi, b_2\zeta) \ll (1 + |\zeta|)^{-2}$, as required.

The right hand side of (8.11) is an integrable function on $\mathbb{R}^2$, and it is immediate that its integral over $\max(|\xi|, |\zeta|) \geq W$ is $O(W^{-1})$. It follows that

$$(8.12) \qquad \mathscr{J} = \mathscr{J}_{\mathbf{a,b}}^+ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \prod_{i=1}^{s} v_0(a_i\xi, b_i\zeta) \, d\zeta \, d\xi$$

exists and that

$$(8.13) \qquad \mathscr{J}(W) = \mathscr{J} + O(W^{-1}).$$

The treatment of the singular series also involves an unconventional element. We shall prove that

$$(8.14) \qquad T(q) \ll q^{2-s/k+\varepsilon}.$$

The difficulty is implied by zero values of b_i because then a factor $S(q, aa_i, 0)$ is present in (8.6), and when $(b; q) = 1$ but $a = q$, this factor is q and will not contribute to the savings needed to prove (8.14). As we shall see momentarily, some other factor in (8.6) will vanish whenever $S(q, aa_i, 0)$ is unduly large. To make this precise, we first apply the method leading from (3.26) to (3.27) to confirm $T(q)$ is multiplicative. Now let p be a prime, $l \geq 1$ and suppose that $p \mid a$ but $p \nmid b$. Since $\mathbf{b}$ is primitive, there is some b_i with $p \nmid b_i$, whence Lemma 2 gives $S(p^l, aa_i, bb_i) = 0$. By (8.6), it follows that

$$T(p^l) = \sum_{\substack{a=1 \\ p \nmid a}}^{p^l} \sum_{b=1}^{p^l} p^{-ls} \prod_{i=1}^{s} S(p^l, aa_i, bb_i)$$

because the remaining pairs a, b with $(a; b; p) = 1$ have $p \mid a$, and hence $p \nmid b$, and then the summand vanishes. For $p \nmid a$, one may apply (2.2) to see that $S(p^l, aa_i, bb_i) \ll p^{l(1-1/k)+\varepsilon}$ (even when $b_i = 0$). This confirms (8.14) for prime powers, and by multiplicativity, (8.14) holds for all q. This estimate shows that the singular series

$$(8.15) \qquad \mathfrak{S} = \mathfrak{S}_{\mathbf{a,b}}^+ = \sum_{q=1}^{\infty} T(q)$$

converges absolutely, and that

$$\sum_{q \leq W} T(q) = \mathfrak{S} + O(W^{\varepsilon-1/k}).$$

We may now combine this with (8.7), (8.8) and (8.13) to conclude as follows.

Lemma 11. *Let $s \geq 3k + 1$ and $r \geq 3$. Suppose that $\mathbf{b}$ is primitive. Then*

$$\iint_{\mathfrak{W}} \mathscr{F}(\alpha, \beta) \, d\alpha \, d\beta = \mathscr{J}_{\mathbf{a,b}}^+ \mathfrak{S}_{\mathbf{a,b}}^+ P^{s-k-1} + O(P^{s-k-1-1/(9k)}).$$

It remains to consider the complementary set of the major arcs $\mathfrak{W}$. Let $\mathfrak{N}$ be the union of the intervals $\{\alpha \in [0,1] : |q\alpha - a| \leq P^{4/5-k}\}$ with $0 \leq a \leq q$, $(a; q) = 1$ and $1 \leq q \leq P^{4/5}$, and let $\mathfrak{n} = [0,1] \setminus \mathfrak{N}$. Now define $\mathfrak{V} = \mathfrak{N} \times [0,1]$, $\mathfrak{v} = \mathfrak{n} \times [0,1]$.

Note that $[0,1]^2$ is the disjoint union of $\mathfrak{W}$, $\mathfrak{V} \setminus \mathfrak{W}$ and $\mathfrak{v}$. For $\mathfrak{w} \in \{\mathfrak{v}, \mathfrak{V} \setminus \mathfrak{W}\}$, and $1 \le j \le 2$, $3 \le i \le s$, let

$$Y_{ij}(\mathfrak{w}) = \iint_{\mathfrak{w}} |f(a_j\alpha, b_j\beta)|^2 |f(a_i\alpha, b_i\beta)|^{s-2} \, d\alpha \, d\beta.$$

Then, applying the simple inequality

$$|f(a_1\alpha, b_1\beta) f(a_2\alpha, b_2\beta)| \le |f(a_1\alpha, b_1\beta)|^2 + |f(a_2\alpha, b_2\beta)|^2$$

together with Hölder's inequality, one infers that

$$(8.16) \qquad \iint_{\mathfrak{w}} |\mathscr{F}(\alpha, \beta)| \, d\alpha \, d\beta \le \sum_{j=1}^{2} \prod_{i=3}^{s} Y_{ij}(\mathfrak{w})^{1/(s-2)}.$$

We proceed by estimating $Y_{ij}(\mathfrak{w})$. First consider the case where $\mathfrak{w} = \mathfrak{v}$, and $3 \le i \le s$ is an index with $b_i = 0$. Then, by orthogonality, and recalling that $b_1 b_2 \ne 0$, one finds that

(8.17)

$$Y_{ij}(\mathfrak{v}) = \int_{\mathfrak{n}} |f(a_i\alpha, 0)|^{s-2} \int_0^1 |f(a_j\alpha, b_j\beta)|^2 \, d\beta \, d\alpha = [P] \int_{\mathfrak{n}} |f(a_i\alpha, 0)|^{s-2} \, d\alpha.$$

We now substitute $\alpha' = \alpha|a_i|$. The definition of $\mathfrak{m}$ in the preamble to Theorem 2 shows that $\{\alpha \colon \alpha/|a_i| \in \mathfrak{n}\} \cap [0,1] \subset \mathfrak{m}$. Since $|f(\alpha, 0)|$ is of period 1, it follows that

$$\int_{\mathfrak{n}} |f(a_i\alpha, 0)|^{s-2} \, d\alpha = \frac{1}{|a_i|} \int_{\{\alpha' \colon \alpha'/|a_i| \in \mathfrak{n}\}} |f(\alpha', 0)|^{s-2} \, d\alpha'$$

$$\le \int_{\mathfrak{m}} |f(\alpha, 0)|^{s-2} \, d\alpha \ll P^{s-2-k} L^{-2}.$$

In the last step, we have used $s \ge 2^k + 2$ and (1.5). By (8.17), this yields

$$(8.18) \qquad Y_{ij}(\mathfrak{v}) \ll P^{s-1-k} L^{-2}.$$

Next consider the case where $\mathfrak{w} = \mathfrak{v}$ and $b_i \ne 0$. Then $b_i b_j \ne 0$, and Hölder's inequality gives

$$(8.19) \quad Y_{ij}(\mathfrak{v}) \le \left(\iint_{\mathfrak{v}} |f(a_j\alpha, b_j\beta)|^s \, d\alpha \, d\beta \right)^{2/s} \left(\iint_{\mathfrak{v}} |f(a_i\alpha, b_i\beta)|^s \, d\alpha \, d\beta \right)^{1-(2/s)}.$$

Here, both integrals are of the same type. A substitution argument similar to the one in the preceding case yields

$$\iint_{\mathfrak{v}} |f(a_j\alpha, b_j\beta)|^s \, d\alpha \, d\beta = \frac{1}{|a_j b_j|} \int_{\{\alpha \colon \alpha/|a_i| \in \mathfrak{n}\}} \int_0^{|b_j|} |f(\alpha, \beta)|^s \, d\beta \, d\alpha$$

$$\le \int_{\mathfrak{m}} \int_0^1 |f(\alpha, \beta)|^s \, d\beta \, d\alpha \ll P^{s-1-k} L^{-2}.$$

Here, we used Theorem 2 for the last inequality. Because the same argument applies with i in place of j, we conclude that (8.18) holds in this case as well. Consequently, by (8.16),

$$(8.20) \qquad \iint_{\mathfrak{v}} |\mathscr{F}(\alpha, \beta)| \, d\alpha \, d\beta \ll P^{s-1-k} L^{-2}.$$

This leaves the set $\mathfrak{V} \setminus \mathfrak{W} = \mathfrak{U}$, say. A successful estimation is possible with the aid of Lemma 10. An inspection of the definitions of the sets $\mathfrak{V}, \mathfrak{W}, \mathfrak{K}$ and $\mathfrak{L}$ shows that whenever $a_j b_j \neq 0$, that

$$\{(\alpha, \beta) \colon (\alpha/|a_j|, \beta/|b_j|) \in \mathfrak{U}\} \cap [0,1]^2 \subset \mathfrak{K} \setminus \mathfrak{L}.$$

Hence, whenever $b_j \neq 0$, one notes that $f(\alpha, \beta)$ is $\mathbb{Z}^2$-periodic to conclude that

$$\iint_{\mathfrak{U}} |f(a_j\alpha, b_j\beta)|^s \, \mathrm{d}\alpha \, \mathrm{d}\beta = \frac{1}{|a_j b_j|} \iint_{\{(\alpha/|a_j|, \beta/|b_j|) \in \mathfrak{U}\}} |f(\alpha, \beta)|^s \, \mathrm{d}\alpha \, \mathrm{d}\beta$$
$$\leq \iint_{\mathfrak{K} \setminus \mathfrak{L}} |f(\alpha, \beta)|^s \, \mathrm{d}\alpha \, \mathrm{d}\beta,$$

and Lemma 10 estimates the last integral as $O(P^{s-k-1-\delta})$ where δ is any real number not exceeding $1/(9k)$. By using Hölder's inequality as in (8.19), it now follows that whenever $j = 1, 2$ and $3 \leq j \leq r$, then

$$(8.21) \qquad\qquad Y_{ij}(\mathfrak{U}) \ll P^{s-k-1-\delta}.$$

For $r < i \leq s$, we merely use orthogonality, and use (1.5) together with a straightforward major arc estimate to infer that

$$Y_{ij}(\mathfrak{U}) \leq \int_0^1 |f(a_i\alpha, 0)|^{s-2} \int_0^1 |f(a_j\alpha, b_j\beta)|^2 \, \mathrm{d}\beta \, \mathrm{d}\alpha$$
$$\leq [P] \int_0^1 |f(a_i\alpha, 0)|^{s-2} \, \mathrm{d}\alpha \ll P^{s-k-1}.$$

We now take $\mathfrak{w} = \mathfrak{U}$ in (8.16). Since $r \geq 3$, the bound (8.21) applies to at least one of the factors, so that one may deduce that

$$(8.22) \qquad\qquad \iint_{\mathfrak{U}} |\mathcal{F}(\alpha, \beta)| \, \mathrm{d}\alpha \, \mathrm{d}\beta \ll P^{s-k-1-\delta/s}.$$

An asymptotic formula for $N_{\mathbf{a},\mathbf{b}}^+(P)$ is now available by combining the conclusions in (8.20) and (8.22) with those in Lemma 11. We summarise our work so far in the next lemma.

Lemma 12. *Let $s \geq 2^k + 2$, $r \geq 3$ and $\mathbf{b}$ be primitive. Then*

$$N_{\mathbf{a},\mathbf{b}}^+(P) = \mathfrak{J}_{\mathbf{a},\mathbf{b}}^+ \mathfrak{S}_{\mathbf{a},\mathbf{b}}^+ P^{s-k-1} + O\big(P^{s-k-1} L^{-2}\big).$$

We are ready to establish Theorem 1 in all cases where $r \geq 3$. Let $N^{(0)}(P)$ be the number of solutions counted by $N_{\mathbf{a},\mathbf{b}}(P)$ where $x_i = 0$ for at least one index i with $1 \leq i \leq s$. For the remaining solutions, we have $x_i \neq 0$ for all i, and we write $\eta_i = x_i/|x_i|$. We group the solutions according to a given value of $\boldsymbol{\eta} \in \{1, -1\}^s$, and substitute $x_i' = \eta_i x_i$ in (1.1). We are then reduced to count solutions in positive integers of the system (1.1) with b_i replaced by $\eta_i b_i$ and a_i replaced by $\eta_i^k a_i$. Thus, on writing $\boldsymbol{\eta}\mathbf{b} = (\eta_i b_i)_{1 \leq i \leq s}$, $\boldsymbol{\eta}\mathbf{a} = (\eta_i^k a_i)_{1 \leq i \leq s}$, we find that

$$(8.23) \qquad\qquad N_{\mathbf{a},\mathbf{b}}(P) = N^{(0)}(P) + \sum_{\eta_i \in \{\pm 1\}} N_{\boldsymbol{\eta}\mathbf{a}, \boldsymbol{\eta}\mathbf{b}}^+(P).$$

We shall establish the estimate

$$(8.24) \qquad\qquad N^{(0)}(P) \ll P^{s-k-1-\delta}$$

at the end of this section. Taking this for granted, the asymptotic formula (1.3) is now available from (8.23) and Lemma 12, with

$$(8.25) \qquad \mathfrak{C}(\mathbf{a}, \mathbf{b}) = \sum_{\eta_i \in \{\pm 1\}} \mathfrak{S}^+_{\boldsymbol{\eta}\mathbf{a}, \boldsymbol{\eta}\mathbf{b}} \mathfrak{J}^+_{\boldsymbol{\eta}\mathbf{a}, \boldsymbol{\eta}\mathbf{b}}.$$

However, substituting $-x$ for x in (2.1) yields $S(q, -a, -b) = S(q, a, b)$ when k is odd, and $S(q, a, -b) = S(q, a, b)$ when k is even. By (8.6) and (8.15), this implies $\mathfrak{S}^+_{\boldsymbol{\eta}\mathbf{a}, \boldsymbol{\eta}\mathbf{b}} = \mathfrak{S}^+_{\mathbf{a}, \mathbf{b}} = \mathfrak{S}$. Moreover, since $T(q)$ is multiplicative, we have

$$(8.26) \qquad \mathfrak{S} = \prod_p \chi_p.$$

By (8.6) and orthogonality

$$(8.27) \qquad \chi_p = \lim_{H \to \infty} p^{-Hs} \sum_{a=1}^{p^H} \sum_{b=1}^{p^H} \prod_{i=1}^{s} S(p^H, aa_i, bb_i) = \lim_{H \to \infty} p^{H(2-s)} M_{\mathbf{a}, \mathbf{b}}(p^H)$$

where $M_{\mathbf{a}, \mathbf{b}}(p^H)$ is the number of incongruent solutions to the system of congruences

$$a_1 x_1^k + \cdots + a_s x_s^k \equiv b_1 x_1 + \cdots + b_s x_s \equiv 0 \bmod p^H.$$

In particular, it follows that χ_p is real and non-negative. Because the product (8.26) converges absolutely, it now follows that there is a $p_0 = p_0(\mathbf{a}, \mathbf{b})$ such that

$$\mathfrak{S} \geq \tfrac{1}{2} \prod_{p \leq p_0} \chi_p.$$

Moreover, the method of proof of Parsell [23, Lemma 7.5] combines with (8.27) to show that whenever (1.1) has a non-singular solution in $\mathbb{Q}_p$, then $\chi_p > 0$. Hence, if (1.1) has non-singular solutions in $\mathbb{Q}_p$ for all primes p, then $\mathfrak{S} > 0$.

We may now conclude that $\mathfrak{C}_{\mathbf{a}, \mathbf{b}} = \mathfrak{S}\mathfrak{J}$ where

$$(8.28) \qquad \mathfrak{J} = \sum_{\eta_i \in \{\pm 1\}} \mathfrak{J}^+_{\boldsymbol{\eta}\mathbf{a}, \boldsymbol{\eta}\mathbf{b}}.$$

Also, provided only that (1.1) has non-singular solutions in all $\mathbb{Q}_p$, we may apply Lemma 12 together with $\mathfrak{S} = \mathfrak{S}^+_{\boldsymbol{\eta}\mathbf{a}, \boldsymbol{\eta}\mathbf{b}}$, $\mathfrak{S} > 0$ and $N^+_{\boldsymbol{\eta}\mathbf{a}, \boldsymbol{\eta}\mathbf{b}}(P) \geq 0$ to conclude that $\mathfrak{J}^+_{\boldsymbol{\eta}\mathbf{a}, \boldsymbol{\eta}\mathbf{b}}$ is real and non-negative, and in view of (8.28), the same is true for $\mathfrak{J}$.

Now suppose that the equations

$$(8.29) \qquad a_1 \xi_1^k + \ldots a_s \xi_s^k = b_1 \xi_1 + \cdots + b_s \xi_s = 0$$

have a non-singular real solution $\boldsymbol{\xi}_0$. By the implicit function theorem, the equations (8.29) define a $(s-2)$-dimensional manifold in a neighbourhood of $\boldsymbol{\xi}_0$. Consequently, we may choose $\boldsymbol{\xi}_0$ such that $\xi_i^{(0)} \neq 0$ for all $1 \leq i \leq s$. By homogeneity, we may also suppose that $|\xi_i^{(0)}| < 1$, for $1 \leq i \leq s$. Now define $\eta_i = \xi_i^{(0)}/|\xi_i^{(0)}|$. Then, the proof of [23, Lemma 7.4] confirms that $\mathfrak{J}^+_{\boldsymbol{\eta}\mathbf{a}, \boldsymbol{\eta}\mathbf{b}} > 0$. Note that once this is established, it follows from (8.28) and the discussion following that formula that $\mathfrak{J} > 0$. Then, recalling the properties of $\mathfrak{S}$ and $\mathfrak{C}_{\mathbf{a}, \mathbf{b}} = \mathfrak{S}\mathfrak{J}$, it follows that $\mathfrak{C}_{\mathbf{a}, \mathbf{b}} > 0$ certainly holds whenever (1.1) admits non-singular solutions in all completions of $\mathbb{Q}$. This establishes Theorem 1.

We are left with the task to prove (8.24). First consider the contribution to $N^{(0)}(P)$ that stems from solutions where two or more of the x_i are zero. If, say, $u \geq 2$ and $x_1 = x_2 = \cdots = x_u = 0$, $x_i \neq 0$ for $u < i \leq s$, then these solutions of

(1.1) satisfy $a_{u+1}x_{u+1}^k + \cdots + a_s x_s^k = 0$, and by writing their number as an integral suitable for application of the Hardy-Littlewood method, the inequalities of Hölder and Hua show that the number of these solutions does not exceed

$$\int_0^1 |f(\alpha, 0)|^{s-u}\,\mathrm{d}\alpha \ll P^{s-2-k+\varepsilon}.$$

By symmetry, this argument applies when any set of u variables x_i vanishes.

If exactly one of the variables vanishes, say $x_1 = 0$, then an analysis of the signs of the other variables similar to (8.23) reduces the problem of counting solutions to (1.1) with $x_1 = 0$ but $1 \le |x_i| \le P$ ($2 \le i \le s$) to an estimate for

$$\int_0^1 \int_0^1 \prod_{i=2}^s |f(a_i\alpha, b_i\beta)|\,\mathrm{d}\alpha\,\mathrm{d}\beta.$$

Note that $b_2 b_3 \ne 0$ (recall $r \ge 3$). An argument similar to (8.16) bounds the above integral by

$$\sum_{j=2}^3 \prod_{i=4}^s \Big(\int_0^1 \int_0^1 |f(a_j\alpha, b_j\beta)|^2 |f(a_i\alpha, b_i\beta)|^{s-3}\,\mathrm{d}\alpha\,\mathrm{d}\beta \Big)^{1/(s-3)}.$$

When $b_i = 0$, we use orthogonality and the classical lemma of Hua to deduce that

$$\int_0^1 \int_0^1 |f(a_j\alpha, b_j\beta)|^2 |f(a_i\alpha, b_i\beta)|^{s-3}\,\mathrm{d}\alpha\,\mathrm{d}\beta \le P \int_0^1 |f(a_i\alpha, 0)|^{s-3}\,\mathrm{d}\alpha \ll P^{s-k-1-\delta}$$

where $\delta > 0$. When $b_i \ne 0$, one may apply Hölder's inequality again, to separate $f(a_j\alpha, b_j\beta)$ from $f(a_i\alpha, b_i\beta)$. An obvious substitution then reduces the problem to that of estimating

$$\int_0^1 \int_0^1 |f(\alpha, \beta)|^{s-1}\,\mathrm{d}\alpha\,\mathrm{d}\beta.$$

Lemma 5 coupled with Hölder's inequality shows that this is $O(P^{s-k-1-\delta})$, for some $\delta > 0$. Hence, the contribution to $N^{(0)}(P)$ from solutions with $x_1 = 0$, $x_i \ne 0$ ($2 \le i \le s$) is acceptable. The argument applies when some other variable vanishes (note that two b_j with $b_j \ne 0$ remain active, which is crucial), and this completes the proof of (8.24).

References

[1] Birch, B. J. Forms in many variables. Proc. Roy. Soc. Ser. A 265 (1961/1962), 245–263.

[2] Boklan, K. D. A reduction technique in Waring's problem. I. Acta Arith. 65 (1993), 147–161.

[3] R. de la Bretèche, Répartition des points rationnels sur la cubique de Segre, Proc. Lond. Math. Soc. (3) 95 (2007), 69–155.

[4] T.D. Browning and D.R. Heath-Brown, Rational points on quartic hypersurfaces. J. Reine Angew. Math. 629 (2009), 37–88.

[5] J. Brüdern, A problem in additive number theory. Math. Proc. Cambridge Philos. Soc. 103 (1988), 27–33.

[6] J. Brüdern, and R.J. Cook, On simultaneous diagonal equations and inequalities. Acta Arith. 62 (1992), 125–149.

[7] J.W.S. Cassels and M.J.T. Guy, On the Hasse principle for cubic surfaces. Mathematika 13 (1966), 111–120.

[8] J.H.H. Chalk, On Hua's estimates for exponential sums. Mathematika 34 (1987), 115–123.

[9] H. Davenport and D.J. Lewis, Cubic equations of additive type. Philos. Trans. Roy. Soc. London Ser. A 261 (1966), 97–136.

[10] H. Davenport and D.J. Lewis, Simultaneous equations of additive type. Philos. Trans. Roy. Soc. London Ser. A 264 (1969), 557–595.

[11] G. Greaves, Some Diophantine equations with almost all solutions trivial. Mathematika 44 (1997), 14–36.

[12] H. Halberstam and H.E. Richert, Sieve methods. London Mathematical Society Monographs, No. 4. Academic Press, London-New York, 1974.

[13] R.R. Hall and G. Tenenbaum, Divisors. Cambridge Tracts in Mathematics, 90. Cambridge University Press, Cambridge, 1988.

[14] M. Harvey, Minor arc moments of Weyl sums. Glasgow Math. J. 55 (2013), 97–113.

[15] C. Hooley, On the representation of a number as the sum of two h-th powers. Math. Z. 84 (1964), 126–136.

[16] C. Hooley, On a new technique and its applications to the theory of numbers. Proc. London Math. Soc. (3) 38 (1979), 115–151.

[17] C. Hooley, On another sieve method and the numbers that are a sum of two hth powers. Proc. London Math. Soc. (3) 43 (1981), 73–109.

[18] C. Hooley, On nonary cubic forms. J. Reine Angew. Math. 386 (1988), 32–98.

[19] C. Hooley, On nonary cubic forms. II. J. Reine Angew. Math. 415 (1991), 95–165.

[20] C. Hooley, On nonary cubic forms. III. J. Reine Angew. Math. 456 (1994), 53–63.

[21] C. Hooley, On another sieve method and the numbers that are a sum of two hth powers. II. J. Reine Angew. Math. 475 (1996), 55–75.

[22] L.K. Hua, Additive theory of prime numbers. Translations of Mathematical Monographs, Vol. 13 American Mathematical Society, Providence, R.I. 1965

[23] S.T. Parsell, Pairs of additive equations of small degree. Acta Arith. 104 (2002), 345–402.

[24] C. Skinner and T.D. Wooley, Sums of two kth powers. J. Reine Angew. Math. 462 (1995), 57–68.

[25] E.C. Titchmarsh, The theory of the Riemann zeta-function. Second edition. Edited and with a preface by D. R. Heath-Brown. The Clarendon Press, Oxford University Press, New York, 1986.

[26] R.C. Vaughan, On Waring's problem for cubes. J. Reine Angew. Math. 365 (1986), 122–170.

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

[28] R.C. Vaughan, A new iterative method in Waring's problem. Acta Math. 162 (1989), 1–71.

[29] R. C. Vaughan, The Hardy-Littlewood Method, *Second Edition*, Cambridge University Press (1997)

[30] R.C. Vaughan, On generating functions in additive number theory. I. Analytic number theory, pp. 436–448, Cambridge Univ. Press, Cambridge, 2009.

[31] R.C. Vaughan and T.D. Wooley, On a certain nonary cubic form and related equations, *Duke Math. J.* 80 (1995), 669–735.

[32] T.D. Wooley, On simultaneous additive equations. II. J. Reine Angew. Math. 419 (1991), 141–198.

[33] T.D. Wooley, On simultaneous additive equations. III. Mathematika 37 (1990), 85–96.

[34] T.D. Wooley, Sums of two cubes. Internat. Math. Res. Notices 1995, 181–184.

[35] T.D. Wooley, Vinogradov's mean value theorem via efficient congruencing. Ann. of Math. (2) 175 (2012), 1575–1627.

[36] T.D. Wooley, The asymptotic formula in Waring's problem. Internat. Math. Res. Notices 2012, 1485-1504.

[37] T.D. Wooley, Vinogradov's mean value theorem via efficient congruencing. II. Duke Math. J. 162 (2013), 673–730

Mathematisches Institut, Bunsenstrasse 3–5, D-37073 Göttingen, Germany
E-mail address: bruedern@uni-math.gwdg.de

Université de Lyon, and Université de Saint-Etienne, Institut Camille Jordan CNRS UMR 5208, 23, rue du Dr P. Michelon, F-42000, Saint-Etienne, France
E-mail address: olivier.robert@univ-st-etienne.fr