ASYMPTOTIC FORMULAE FOR PAIRS OF DIAGONAL EQUATIONS

Jörg Brüdern and Trevor D. Wooley*

1. Introduction. Consider a system of diagonal equations

$$\sum_{j=1}^{s} a_{ij}x_j^k = 0 \quad (1 \leqslant i \leqslant r), \tag{1.1}$$

satisfying the property that the (fixed) integral coefficient matrix (a_{ij}) contains no singular $r \times r$ submatrix. A recent note of the authors [3] establishes that whenever $k \geqslant 3$ and $s > (3r+1)2^{k-2}$, then the expected asymptotic formula holds for the number $N(P)$ of integral solutions $\mathbf{x}$ of (1.1) with $|x_i| \leqslant P$ $(1 \leqslant i \leqslant s)$. To be precise, as $P \to \infty$ one has

$$N(P) = v_\infty \Bigl(\prod_p v_p\Bigr) P^{s-rk} + o(P^{s-rk}), \tag{1.2}$$

where v_∞ is the area of the manifold defined by (1.1) in the box $[-1,1]^s$, and

$$v_p = \lim_{h\to\infty} p^{h(r-s)} \text{card}\Bigl\{\mathbf{x} \in (\mathbb{Z}/p^h\mathbb{Z})^s : \sum_{j=1}^{s} a_{ij}x_j^k \equiv 0 \pmod{p^h}\ (1 \leqslant i \leqslant r)\Bigr\}.$$

In this note we concentrate on the situation wherein $r = 2$, providing an alternate derivation of this conclusion in which weaker conditions are imposed on the coefficient matrix. Moreover, when $k \geqslant 6$, the constraint on the number s of variables is also relaxed.

Henceforth we fix r to be 2, and we rewrite the system (1.1) in the form

$$a_1x_1^k + \dots + a_sx_s^k = b_1x_1^k + \dots + b_sx_s^k = 0, \tag{1.3}$$

drawing the obvious correspondence between the coefficient matrix (a_{ij}) and the s-tuples $\mathbf{a}, \mathbf{b}$. The ordered pairs (a_i, b_i) $(1 \leqslant i \leqslant s)$ determine s points on the projective line, of which l, say, are distinct. We may suppose that the respective multiplicities $m_1, \dots, m_l$ of these points satisfy $m_1 \geqslant m_2 \geqslant \dots \geqslant m_l \geqslant 1$ and $m_1 + \dots + m_l = s$. It is convenient to refer to the l-tuple $(m_1, \dots, m_l)$ as the *profile* of the pair of equations (1.3).

Theorem 1.1. *Let the profile of the system* (1.3) *be* $(m_1, \dots, m_l)$.

(i) Suppose that $k \geqslant 3$ *and* $s > \frac{7}{4}2^k$*, and that the profile satisfies* $m_1 \leqslant s - 2^k$ *and* $m_1 + m_2 \leqslant s - 2^{k-2}$*. Then the asymptotic formula* (1.2) *holds.*

(ii) Suppose that $k \geqslant 6$ *and* $s > \frac{13}{8}2^k$*, and that the profile satisfies* $m_1 \leqslant s - \frac{15}{16}2^k$ *and* $m_1 + m_2 \leqslant s - 2^{k-2}$*. Then the asymptotic formula* (1.2) *holds.*

The first conclusion of Theorem 1.1 has the same strength as the principal conclusion of [3], although the hypotheses on the profile of the pair of equations are now explicitly weaker than are required in [3]. The second conclusion of Theorem 1.1, on the other hand, represents an improvement on this earlier work in so far as the condition $s > \frac{7}{4}2^k$ is now replaced by $s > \frac{13}{8}2^k$ for $k \geqslant 6$. When $k \geqslant 9$ or thereabouts, Vinogradov's methods may be applied to provide sharper bounds still. These results

2000 *Mathematics Subject Classification.* 11D72, 11P55.

Key words and phrases. Exponential sums, diophantine equations.

*The authors are grateful to the Max Planck Institute in Bonn for its generous hospitality during the period in which this paper was written. The second author was also supported in part by NSF grant DMS-0140523

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

should also be compared with the earlier work of Cook [5], in the refined form that follows from an educated application of Vaughan's mean value estimates [9], [10] associated with the asymptotic formula in Waring's problem. This work shows that whenever $k \geqslant 3$ and $s \geqslant 2^{k+1}$, and the profile $(m_1, \ldots, m_l)$ of the system (1.3) satisfies the condition $m_1 \leqslant s - 2^k$, then the asymptotic formula (1.2) holds. By considering equations containing disjoint sets of variables, it is apparent that the latter hypothesis on m_1 cannot be relaxed without first establishing the validity of the expected asymptotic formula for a single diagonal equation in fewer than 2^k variables. For the moment, at least, the latter objective remains beyond reach for $k \leqslant 5$.

Our strategy for proving Theorem 1.1 involves relating a mean value of exponential sums with a mean square of certain Fourier coefficients, wherein the summation is over a potentially restricted set. This method is related to that occurring in recent work of the authors on paucity problems (see [4]). The principal conclusion of this discussion is recorded in Theorem 2.2 below. The application of this new mean value estimate to establish Theorem 1.1 is routine, and summarily executed in §3.

Throughout, the letter ε will denote a sufficiently small positive number, and P will be a large real number. We use $\ll$ and $\gg$ to denote Vinogradov's notation. In an effort to simplify our account, whenever ε appears in a statement, we assert that the statement holds for every positive number ε. The "value" of ε may consequently change from statement to statement.

2. An auxiliary mean value estimate. The purpose of this section is to provide an estimate for a certain mean value of exponential sums that may be thought of as a restricted mean square of Fourier coefficients. When P is a large positive number, put

$$f(\gamma) = \sum_{|x| \leqslant P} e(\gamma x^k),$$

where, as usual, we write $e(z)$ for $e^{2\pi i z}$. We consider a non-negative function F in $L^2([0,1))$, extended in the natural way to a periodic function on $\mathbb{R}$ with period 1, and satisfying the condition that $F(\gamma) = F(-\gamma)$ for each $\gamma \in \mathbb{R}$. When $\mathfrak{B} \subseteq [0,1)$, we then write

$$R(n; \mathfrak{B}) = \int_{\mathfrak{B}} F(\gamma) e(-n\gamma)\, \mathrm{d}\gamma,$$

and we abbreviate $R(n; [0,1))$ simply to $R(n)$. Also, when a is a non-zero integer, we denote by $\rho_a(n)$ the number of representations of the integer n in the shape

$$a \sum_{i=1}^{2^{k-3}} (x_i^k - y_i^k) = n,$$

with $|x_i|, |y_i| \leqslant P$ $(1 \leqslant i \leqslant 2^{k-3})$.

Lemma 2.1. *Suppose that a is a non-zero integer. Then whenever $k \geqslant 3$ and $\mathfrak{B} \subseteq [0,1)$, one has*

$$\sum_{n \in \mathbb{Z}} \rho_a(n) |R(n; \mathfrak{B})|^2 \ll P^{2^{k-2}-1} \left(\int_{\mathfrak{B}} F(\gamma)\, \mathrm{d}\gamma \right)^2 + P^{2^{k-2}-k+1+\varepsilon} \int_{\mathfrak{B}} F(\gamma)^2\, \mathrm{d}\gamma. \tag{2.1}$$

Proof. We begin by noting that

$$\begin{aligned}\sum_{n \in \mathbb{Z}} \rho_a(n) |R(n; \mathfrak{B})|^2 &= \sum_{n \in \mathbb{Z}} \rho_a(n) \int_{\mathfrak{B}} \int_{\mathfrak{B}} F(\alpha) F(-\beta) e(n(\beta - \alpha))\, \mathrm{d}\alpha\, \mathrm{d}\beta \\ &= \int_{\mathfrak{B}} \int_{\mathfrak{B}} F(\alpha) F(-\beta) |f(a(\beta - \alpha))|^{2^{k-2}}\, \mathrm{d}\alpha\, \mathrm{d}\beta. \end{aligned} \tag{2.2}$$

Next, as a consequence of the Weyl differencing lemma (see, for example, Lemma 2.3 of [11]), one has

$$|f(\xi)|^{2^{k-2}} \ll P^{2^{k-2}-1} + P^{2^{k-2}-k+1} \sum_{0 < |h| \leqslant k!(2P)^k} c_h e(h\xi),$$

where the integers c_h satisfy $c_h = O(|h|^\varepsilon)$. On substituting this upper bound into the right hand side of (2.2), we find that

$$\sum_{n\in\mathbb{Z}} \rho_a(n)|R(n;\mathfrak{B})|^2 \ll P^{2^{k-2}-1}\left(\int_{\mathfrak{B}} F(\gamma)\,\mathrm{d}\gamma\right)^2 + P^{2^{k-2}-k+1} I, \tag{2.3}$$

where

$$I = \sum_{0<|h|\leqslant k!(2P)^k} c_h \int_{\mathfrak{B}}\int_{\mathfrak{B}} F(\alpha)F(-\beta)e(a(\beta-\alpha)h)\,\mathrm{d}\alpha\,\mathrm{d}\beta.$$

But in view of the above-cited bound for c_h, one has

$$I \ll \sum_{0<|h|\leqslant k!(2P)^k} |h|^\varepsilon |R(ah;\mathfrak{B})|^2.$$

Then it follows from Bessel's inequality that

$$I \ll P^\varepsilon \sum_{n\in\mathbb{Z}} |R(n;\mathfrak{B})|^2 \leqslant P^\varepsilon \int_{\mathfrak{B}} F(\gamma)^2\,\mathrm{d}\gamma.$$

The conclusion (2.1) of the lemma now follows on recalling (2.3).

We extract from Lemma 2.1 the two consequences recorded in the following theorem. In this context, we write $\Lambda_i = a_i\alpha + b_i\beta$ $(1\leqslant i\leqslant s)$.

Theorem 2.2. *Suppose that Λ_u, Λ_v and Λ_w are pairwise linearly independent linear forms in α and β.*

(i) When $k\geqslant 3$, one has

$$\int_0^1\int_0^1 |f(\Lambda_u)^3 f(\Lambda_v)^3 f(\Lambda_w)|^{2^{k-2}}\,\mathrm{d}\alpha\,\mathrm{d}\beta \ll P^{\frac{7}{4}2^k-2k+\varepsilon}.$$

(ii) When $k\geqslant 6$, one has

$$\int_0^1\int_0^1 |f(\Lambda_u)^{11} f(\Lambda_v)^{11} f(\Lambda_w)^4|^{2^{k-4}}\,\mathrm{d}\alpha\,\mathrm{d}\beta \ll P^{\frac{13}{8}2^k-2k+\varepsilon}.$$

Proof. We observe first that on considering the underlying diophantine system, and taking linear combinations of the relevant equations, the respective bounds of the theorem follow whenever they can be established in the special case $\Lambda_u = a\alpha$, $\Lambda_v = b\beta$, $\Lambda_w = c\alpha + d\beta$, for non-zero integers a, b, c, d. We may assume without loss, moreover, that $(a,c)=(b,d)=1$. Let $T_s(n)$ denote the number of representations of the integer n in the form

$$n = \sum_{i=1}^s (x_i^k - y_i^k),$$

with $|x_i|, |y_i| \leqslant P$ $(1\leqslant i\leqslant s)$. Then in the special case under consideration, again making use of the underlying diophantine equations, we find that whenever $r,t\in\mathbb{N}$ one has

$$\int_0^1\int_0^1 |f(\Lambda_u)^r f(\Lambda_v)^r f(\Lambda_w)^t|^2\,\mathrm{d}\alpha\,\mathrm{d}\beta = \Xi,$$

where

$$\Xi = \sum_{\substack{n_1,n_2,n_3\in\mathbb{Z}\\ an_1=cn_3\\ bn_2=dn_3}} T_r(n_1)T_r(n_2)T_t(n_3).$$

On writing $[a,b]$ for the least common multiple of a and b, it therefore follows that

$$\Xi = \sum_{\substack{n\in\mathbb{Z}\\ [a,b]\mid n}} T_r(cn/a)T_r(dn/b)T_t(n).$$

On applying Cauchy's inequality, we may conclude that

$$\Xi \leqslant \Xi(a,c)^{1/2}\Xi(b,d)^{1/2},$$

where we write

$$\Xi(g,h) = \sum_{\substack{n\in\mathbb{Z}\\ g|n}} T_r(hn/g)^2 T_t(n) = \sum_{m\in\mathbb{Z}} T_r(hm)^2 T_t(gm).$$

But when $(g,h)=1$, on considering the underlying diophantine equations, one finds that

$$\Xi(g,h) = \int_0^1\int_0^1 |f(g\alpha)^r f(g\beta)^r f(h(\alpha+\beta))^t|^2 \,\mathrm{d}\alpha\,\mathrm{d}\beta.$$

In view of this discussion, it suffices to establish the special case of the theorem in which $\Lambda_u = g\alpha$, $\Lambda_v = g\beta$, $\Lambda_w = h(\alpha+\beta)$, wherein g and h are non-zero integers with $(g,h)=1$.

Next we define a Hardy-Littlewood dissection. We take $\mathfrak{M}$ to be the union of the intervals

$$\mathfrak{M}(q,a) = \{\alpha\in[0,1) : |q\alpha - a| \leqslant q^{-1}P^{1-k}\},$$

with $0\leqslant a\leqslant q\leqslant P$ and $(a,q)=1$, and put $\mathfrak{m} = [0,1)\setminus\mathfrak{M}$. We then apply Lemma 2.1 with $F(\gamma) = |f(g\gamma)|^{2r}$, for appropriate choices of r, and note that

$$\sum_{n\in\mathbb{Z}} \rho_h(n)|R(n)|^2 \ll \sum_{n\in\mathbb{Z}} \rho_h(n)|R(n;\mathfrak{m})|^2 + \sum_{n\in\mathbb{Z}} \rho_h(n)|R(n;\mathfrak{M})|^2. \tag{2.4}$$

For the first part of the theorem we suppose that $k\geqslant 3$ and take $r = 3\cdot 2^{k-3}$. Here it is useful to recall that, from Hua's lemma (see Lemma 2.5 of [11]) in combination with Schwarz's inequality, one has

$$\int_0^1 |f(g\gamma)|^{2r}\,\mathrm{d}\gamma \leqslant \left(\int_0^1 |f(\gamma)|^{2^k}\,\mathrm{d}\gamma\right)^{1/2}\left(\int_0^1 |f(\gamma)|^{2^{k-1}}\,\mathrm{d}\gamma\right)^{1/2} \ll P^{2r-k+1/2+\varepsilon}. \tag{2.5}$$

Also, again applying Hua's lemma, but now in combination with Weyl's inequality (see Lemma 2.4 of [11]),

$$\begin{aligned}\int_{\mathfrak{m}} |f(g\gamma)|^{4r}\,\mathrm{d}\gamma &\leqslant \left(\sup_{\gamma\in\mathfrak{m}} |f(g\gamma)|\right)^{2^{k-1}} \int_0^1 |f(g\gamma)|^{2^k}\,\mathrm{d}\gamma\\ &\ll P^{\varepsilon}(P^{1-2^{1-k}})^{2^{k-1}} P^{2^k-k} \ll P^{4r-k-1+\varepsilon}.\end{aligned} \tag{2.6}$$

In this way, it follows from Lemma 2.1 that

$$\begin{aligned}\sum_{n\in\mathbb{Z}} \rho_h(n)|R(n;\mathfrak{m})|^2 &\ll P^{2^{k-2}-1}\left(\int_{\mathfrak{m}} |f(g\gamma)|^{2r}\,\mathrm{d}\gamma\right)^2 + P^{2^{k-2}-k+1+\varepsilon}\int_{\mathfrak{m}} |f(g\gamma)|^{4r}\,\mathrm{d}\gamma\\ &\ll P^{4r+2^{k-2}-2k+\varepsilon}.\end{aligned} \tag{2.7}$$

The contribution arising from the major arcs is easily controlled. Indeed, the methods of Chapter 4 of [11] establish that $R(n;\mathfrak{M}) \ll P^{2r-k+\varepsilon}$ whenever $2r\geqslant \max\{4,k+1\}$, as we may assume. Thus we find that

$$\sum_{n\in\mathbb{Z}} \rho_h(n)|R(n;\mathfrak{M})|^2 \ll (P^{2r-k+\varepsilon})^2 \sum_{n\in\mathbb{Z}} \rho_h(n) \ll P^{4r+2^{k-2}-2k+2\varepsilon}. \tag{2.8}$$

On substituting (2.7) and (2.8) into (2.4), we conclude that

$$\sum_{n\in\mathbb{Z}} \rho_h(n)|R(n)|^2 \ll P^{4r+2^{k-2}-2k+\varepsilon},$$

and the desired conclusion in part (i) follows on applying the relation (2.2), with $\mathfrak{B} = [0,1)$, and making a change of variables.

The proof of the second part of the theorem follows in like manner, but makes use of Heath-Brown's mean value estimate

$$\int_0^1 |f(\gamma)|^{\frac{7}{8}2^k}\,\mathrm{d}\gamma \ll P^{\frac{7}{8}2^k-k+\varepsilon} \tag{2.9}$$

(see [8]) in place of Hua's lemma at suitable points of the argument. We now take $r = 11\cdot 2^{k-5}$, and note that by Hua's lemma together with (2.9), an application of Schwarz's inequality yields the estimate

$$\begin{aligned}\int_0^1 |f(g\gamma)|^{2r}\,\mathrm{d}\gamma &\leqslant \left(\int_0^1 |f(\gamma)|^{\frac{7}{8}2^k}\,\mathrm{d}\gamma\right)^{1/2}\left(\int_0^1 |f(\gamma)|^{2^{k-1}}\,\mathrm{d}\gamma\right)^{1/2}\\ &\ll P^{\varepsilon}(P^{\frac{7}{8}2^k-k})^{1/2}(P^{2^{k-1}-k+1})^{1/2}.\end{aligned}$$

Thus the estimate (2.5) remains valid, and one may confirm the upper bound (2.6) in these circumstances by noting that we now have

$$\begin{aligned}\int_{\mathfrak{m}} |f(g\gamma)|^{4r}\,\mathrm{d}\gamma &\leqslant \left(\sup_{\gamma\in\mathfrak{m}}|f(g\gamma)|\right)^{2^{k-1}}\int_0^1 |f(\gamma)|^{\frac{7}{8}2^k}\,\mathrm{d}\gamma\\ &\ll P^{\varepsilon}(P^{1-2^{1-k}})^{2^{k-1}}(P^{\frac{7}{8}2^k-k}).\end{aligned}$$

The upper bounds (2.7) and (2.8) now follow just as in the argument applied to establish the first part of the theorem, and in this way the desired conclusion follows as before.

3. Application of the Hardy-Littlewood method. With the mean value estimates embodied in Theorem 2.2 now at our disposal, the proof of Theorem 1.1 is in principle routine. The only difficulties that arise are generated by the possibility that in the profile $\mathbf{m}$ of the pair of equations (1.3), the parameter m_1 might be unpleasantly large. Before proceeding further, we define the Hardy-Littlewood dissection underpinning our analysis. We take $\delta = 1/100$, and define the set of major arcs $\mathfrak{N}$ to be the union of the boxes

$$\mathfrak{N}(q,a,b) = \{(\alpha,\beta)\in[0,1)^2 \,:\, |q\alpha-a|\leqslant P^{\delta-k},\ |q\beta-b|\leqslant P^{\delta-k}\}, \tag{3.1}$$

with $0\leqslant a,b\leqslant q\leqslant P^{\delta}$ and $(q,a,b)=1$. We then define the corresponding set of minor arcs $\mathfrak{n}$ by putting $\mathfrak{n}=[0,1)^2\setminus\mathfrak{N}$. When $\mathfrak{B}\subseteq[0,1)^2$ is measureable, it is convenient henceforth to write

$$N(P;\mathfrak{B}) = \iint_{\mathfrak{B}} f(\Lambda_1)\dots f(\Lambda_s)\,\mathrm{d}\alpha\,\mathrm{d}\beta.$$

Here we assume implicitly that $s>\frac{7}{4}2^k$ when $3\leqslant k\leqslant 5$, that $s>\frac{13}{8}2^k$ when $k\geqslant 6$, and that the profile of the implicit pair of equations satisfies the conditions of the statement of Theorem 1.1.

We begin by observing that standard methods based on work of Davenport and Lewis [6], [7] (see also Brüdern and Cook [2]) easily establish the asymptotic formula

$$N(P;\mathfrak{N}) = v_\infty\Big(\prod_p v_p\Big)P^{s-2k} + o(P^{s-2k}), \tag{3.2}$$

where the local factors v_ϖ are defined as in the introduction. Since $N(P)=N(P;[0,1)^2)$, it now remains to demonstrate that the contribution of the minor arcs $\mathfrak{n}$ to $N(P)$ is $o(P^{s-2k})$. In order to facilitate our argument at this point, we introduce an auxiliary one dimensional Hardy-Littlewood dissection. We put $\tau=10^{-3}$, and define the auxiliary major arcs $\mathfrak{P}$ to be the union of the intervals

$$\mathfrak{P}(q,a) = \{\alpha\in[0,1) \,:\, |q\alpha-a|\leqslant P^{\tau-k}\},$$

with $0\leqslant a\leqslant q\leqslant P^{\tau}$ and $(a,q)=1$. We also put $\mathfrak{p}=[0,1)\setminus\mathfrak{P}$.

Define the parameter t to be $3\cdot 2^{k-2}$ when $3\leqslant k\leqslant 5$, and to be $11\cdot 2^{k-4}$ when $k\geqslant 6$. Also, write r for 2^{k-2}. Then by making repeated use of the elementary inequality $|z_1\dots z_n|\leqslant |z_1|^n+\dots+|z_n|^n$, one finds that our hypotheses on the profile of the system (1.3) suffice to ensure that

$$|f(\Lambda_1)\dots f(\Lambda_s)| \ll \sum_{u,v,w} |f(\Lambda_u)^{s-t-r}f(\Lambda_v)^t f(\Lambda_w)^r|,$$

where the summation runs over all triples (u,v,w) with $1\leqslant u,v,w\leqslant s$ for which the points (a_j,b_j) $(j=u,v,w)$ are all distinct on $\mathbb{P}^1$. It therefore follows that there is a choice of (u,v,w) satisfying the latter condition for which

$$N(P;\mathfrak{n})\ll \iint_{\mathfrak{n}} |f(\Lambda_u)^{s-t-r}f(\Lambda_v)^t f(\Lambda_w)^r|\,\mathrm{d}\alpha\,\mathrm{d}\beta. \tag{3.3}$$

We estimate this integral by dividing the minor arcs into three sets, namely

$$\mathfrak{n}_1=\{(\alpha,\beta)\in\mathfrak{n}\,:\,\Lambda_u\in\mathfrak{p}\pmod 1\},$$

$$\mathfrak{n}_2=\{(\alpha,\beta)\in\mathfrak{n}\,:\,\Lambda_u\in\mathfrak{P}\pmod 1\text{ and }\Lambda_v\in\mathfrak{p}\pmod 1\},$$

$$\mathfrak{n}_3=\{(\alpha,\beta)\in\mathfrak{n}\,:\,\Lambda_u\in\mathfrak{P}\pmod 1\text{ and }\Lambda_v\in\mathfrak{P}\pmod 1\}.$$

We first consider the contribution of the set $\mathfrak{n}_1$ within (3.3). It follows from Weyl's inequality (see Lemma 2.4 of [11]) that

$$\sup_{(\alpha,\beta)\in\mathfrak{n}_1}|f(\Lambda_u)|\leqslant \sup_{\gamma\in\mathfrak{p}}|f(\gamma)|\ll P^{1-\tau 2^{1-k}+\varepsilon}.$$

Then in view of Theorem 2.2, one has

$$\begin{aligned}N(P;\mathfrak{n}_1)&\leqslant \Bigl(\sup_{(\alpha,\beta)\in\mathfrak{n}_1}|f(\Lambda_u)|\Bigr)^{s-2t-r}\int_0^1\int_0^1 |f(\Lambda_u)^t f(\Lambda_v)^t f(\Lambda_w)^r|\,\mathrm{d}\alpha\,\mathrm{d}\beta\\ &\ll P^{\varepsilon}(P^{1-\tau 2^{1-k}})^{s-2t-r}(P^{2t+r-2k})=o(P^{s-2k}).\end{aligned} \tag{3.4}$$

Next we consider the contribution of the set $\mathfrak{n}_2$ within (3.3). Here we note that the methods of Chapter 4 of [11] show that whenever $\sigma\geqslant k+2$, then one has

$$\int_{\mathfrak{P}}|f(\gamma)|^{\sigma}\,\mathrm{d}\gamma\ll P^{\sigma-k}. \tag{3.5}$$

Also, it follows from the methods underlying Theorem A of Vaughan [10] (when $k=4,5$) and Lemma F of section 4 of Boklan [1] (when $k=3$; see Vaughan [9] for an earlier, slightly weaker conclusion) that for $3\leqslant k\leqslant 5$, one has

$$\int_{\mathfrak{p}}|f(\gamma)|^{r+t}\,\mathrm{d}\gamma\ll P^{r+t-k}(\log P)^{\varepsilon-2}\quad\text{and}\quad \int_0^1|f(\gamma)|^{r+t}\,\mathrm{d}\gamma\ll P^{r+t-k}.$$

When $k\geqslant 6$, on the other hand, the same conclusions follow by combining Weyl's inequality with Heath-Brown's mean value estimate (2.9), and in the case of the second inequality, also the use of (3.5) with $\sigma=r+t$. In any case, therefore, it follows from Hölder's inequality that

$$N(P;\mathfrak{n}_2)\leqslant \Bigl(\sup_{(\alpha,\beta)\in\mathfrak{n}_2}|f(\Lambda_u)|\Bigr)^{s-2t-r}J_v^{t/(r+t)}J_w^{r/(r+t)}, \tag{3.6}$$

where

$$J_z=\iint_{\mathfrak{n}_2}|f(\Lambda_u)^t f(\Lambda_z)^{r+t}|\,\mathrm{d}\alpha\,\mathrm{d}\beta\quad (z=v,w).$$

But when $(\alpha,\beta)\in\mathfrak{n}_2$, one has $\Lambda_u\in\mathfrak{P}\pmod 1$ and $\Lambda_v\in\mathfrak{p}\pmod 1$, and thus a change of variables reveals that

$$J_v\ll\Bigl(\int_{\mathfrak{P}}|f(\gamma_1)|^t\,\mathrm{d}\gamma_1\Bigr)\Bigl(\int_{\mathfrak{p}}|f(\gamma_2)|^{r+t}\,\mathrm{d}\gamma_2\Bigr)\ll (P^{t-k})(P^{r+t-k}(\log P)^{\varepsilon-2}).$$

Meanwhile, in similar fashion, one obtains the bound

$$J_w\ll\Bigl(\int_{\mathfrak{P}}|f(\gamma_1)|^t\,\mathrm{d}\gamma_1\Bigr)\Bigl(\int_0^1|f(\gamma_2)|^{r+t}\,\mathrm{d}\gamma_2\Bigr)\ll (P^{t-k})(P^{r+t-k}).$$

On substituting these estimates into (3.6), therefore, and applying a trivial estimate for $|f(\Lambda_u)|$, we may conclude that

$$N(P;\mathfrak{n}_2) \ll (P^{s-2t-r})(P^{r+2t-2k}(\log P)^{\varepsilon-1}) = o(P^{s-2k}). \tag{3.7}$$

Finally, we consider the contribution of the set $\mathfrak{n}_3$ within (3.3), and it is here that we discover the joker in the pack. Suppose that $(\alpha,\beta)\in\mathfrak{n}_3$, and write (α_1,α_2) for the ordered pair in $[0,1)^2$ for which $(\Lambda_u,\Lambda_v)\equiv(\alpha_1,\alpha_2) \pmod 1$. Then for $i=1,2$, there exist $r_i\in\mathbb{Z}$ and $q_i\in\mathbb{N}$ with

$$0\leqslant r_i\leqslant q_i\leqslant P^{\tau},\quad (r_i,q_i)=1\quad\text{and}\quad |q_i\alpha_i-r_i|\leqslant P^{\tau-k}. \tag{3.8}$$

By taking suitable linear combinations, the final inequality of (3.8) may be employed to isolate α and β. Write

$$q'=(a_ub_v-a_vb_u)q_1q_2,\quad a'=b_vr_1q_2-b_ur_2q_1,\quad b'=a_ur_2q_1-a_vr_1q_2,$$

and then put $\omega=\nu(q',a',b')$, where ν is the sign of q'. Next define $q=q'/\omega$, $a^*=a'/\omega$, $b^*=b'/\omega$, and take (a,b) to be the unique ordered pair in $[0,q)^2$ with $(a^*,b^*)\equiv(a,b) \pmod q$. Then we find in this way that $0\leqslant a,b\leqslant q\ll P^{2\tau}$, $(a,b,q)=1$,

$$|q\alpha-a|\ll P^{2\tau-k}\quad\text{and}\quad |q\beta-b|\ll P^{2\tau-k}.$$

Here, the implicit constants depend at most on (a_j,b_j) for $j=u,v$. But since $2\tau<\delta$, and we may suppose that P is sufficiently large, an inspection of (3.1) forces us to conclude that $(\alpha,\beta)\in\mathfrak{N}$, and this contradicts our assumption that $(\alpha,\beta)\in\mathfrak{n}_3\subseteq\mathfrak{n}$. Thus we see that $\mathfrak{n}_3=\emptyset$, so that necessarily

$$N(P;\mathfrak{n}_3)=0. \tag{3.9}$$

Finally, on combining the estimates (3.4), (3.7) and (3.9), we see that

$$N(P;\mathfrak{n})\leqslant N(P;\mathfrak{n}_1)+N(P;\mathfrak{n}_2)+N(P;\mathfrak{n}_3)=o(P^{s-2k}).$$

In combination with the asymptotic formula (3.2), we thus deduce that

$$N(P)=N(P;\mathfrak{N})+N(P;\mathfrak{n})\sim v_\infty\Big(\prod_p v_p\Big)P^{s-2k}+o(P^{s-2k}),$$

and this completes the proof of Theorem 1.1.

References

1. K. D. Boklan, *A reduction technique in Waring's problem, I*, Acta Arith. **65** (1993), 147–161.
2. J. Brüdern and R. J. Cook, *On simultaneous diagonal equations and inequalities*, Acta Arith. **62** (1992), 125–149.
3. J. Brüdern and T. D. Wooley, *Hua's lemma and simultaneous diagonal equations*, Bull. London Math. Soc. **34** (2002), 279–283.
4. J. Brüdern and T. D. Wooley, *The paucity problem for certain pairs of diagonal equations*, Quart. J. Math. (in press).
5. R. J. Cook, *Pairs of additive equations*, Michigan Math. J. **19** (1972), 325–331.
6. H. Davenport and D. J. Lewis, *Cubic equations of additive type*, Philos. Trans. Roy. Soc. London Ser. A **261** (1966), 97–136.
7. H. Davenport and D. J. Lewis, *Simultaneous equations of additive type*, Philos. Trans. Roy. Soc. London Ser. A **264** (1969), 557–595.
8. D. R. Heath-Brown, *Weyl's inequality, Hua's inequality, and Waring's problem*, J. London Math. Soc. (2) **38** (1988), 216–230.
9. R. C. Vaughan, *On Waring's problem for cubes*, J. Reine Angew. Math. **365** (1986), 122–170.
10. R. C. Vaughan, *On Waring's problem for smaller exponents. II*, Mathematika **33** (1986), 6–22.
11. R. C. Vaughan, *The Hardy-Littlewood method, 2nd. ed.*, Cambridge University Press, Cambridge, 1997.

JB: Mathematisches Institut A, Universität Stuttgart, Pfaffenwaldring 57, D-70511 Stuttgart, Germany

E-mail address: bruedern@mathematik.uni-stuttgart.de

TDW: Department of Mathematics, University of Michigan, East Hall, 525 East University Ave., Ann Arbor, MI 48109-1109, U.S.A.

E-mail address: wooley@umich.edu