THE HASSE PRINCIPLE FOR PAIRS OF DIAGONAL CUBIC FORMS

Jörg Brüdern and Trevor D. Wooley*

Abstract. By means of the Hardy-Littlewood method, we apply a new mean value theorem for exponential sums to confirm the truth, over the rational numbers, of the Hasse principle for pairs of diagonal cubic forms in thirteen or more variables.

1. Introduction. Early work of Lewis [14] and Birch [3, 4], now almost a half-century old, shows that pairs of quite general homogeneous cubic equations possess non-trivial integral solutions whenever the dimension of the corresponding intersection is suitably large (modern refinements have reduced this permissible affine dimension to 826; see [13]). When s is a natural number, let a_j, b_j $(1 \le j \le s)$ be fixed rational integers. Then the pioneering work of Davenport and Lewis [12] employs the circle method to show that the pair of simultaneous *diagonal* cubic equations

$$a_1x_1^3 + a_2x_2^3 + \ldots + a_sx_s^3 = b_1x_1^3 + b_2x_2^3 + \ldots + b_sx_s^3 = 0, \tag{1.1}$$

possess a non-trivial solution $\mathbf{x} \in \mathbb{Z}^s \setminus \{\mathbf{0}\}$ provided only that $s \ge 18$. Their analytic work was simplified by Cook [10] and enhanced by Vaughan [16]; these authors showed that the system (1.1) necessarily possesses non-trivial integral solutions in the cases $s = 17$ and $s = 16$, respectively. Subject to a local solubility hypothesis, a corresponding conclusion was obtained for $s = 15$ by Baker and Brüdern [2], and for $s = 14$ by Brüdern [5]. Our purpose in this paper is the proof of a similar result that realises the sharpest conclusion attainable by any version of the circle method as currently envisioned, even if one were to be equipped with the most powerful mean value estimates for Weyl sums conjectured to hold.

Theorem 1. *Suppose that $s \ge 13$, and that $a_j, b_j \in \mathbb{Z}$ $(1 \le j \le s)$. Then the pair of equations* (1.1) *has a non-trivial solution in rational integers if and only if it has a non-trivial solution in the 7-adic field. In particular, the Hasse principle holds for the system* (1.1) *provided only that $s \ge 13$.*

1991 *Mathematics Subject Classification.* 11D72, 11L07, 11E76, 11P55.

Key words and phrases. Diophantine equations, exponential sums, Hardy-Littlewood method.

*Supported in part by NSF grant DMS-010440. The authors are grateful to the Max Planck Institut in Bonn for its generous hospitality during the period in which this paper was conceived.

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

When $s \geq 13$, the conclusion of Theorem 1 confirms the Hasse principle for the system (1.1) in a particularly strong form: any local obstruction to solubility must necessarily be 7-adic. Similar conclusions follow from the earlier cited work of Baker and Brüdern [2] and Brüdern [5] under the more stringent conditions $s \geq 15$ and $s \geq 14$, respectively.

The conclusion of Theorem 1 is best possible in several respects. First, when $s = 12$, there may be arbitrarily many p-adic obstructions to global solubility. For example, let $\mathcal{S}$ denote any finite set of primes $p \equiv 1 \pmod 3$, and write q for the product of all the primes in $\mathcal{S}$. Choose any number $a \in \mathbb{Z}$ that is a cubic non-residue modulo p for all $p \in \mathcal{S}$, and consider the form

$$\Psi(x_1, \dots, x_6) = (x_1^3 - ax_2^3) + q(x_3^3 - ax_4^3) + q^2(x_5^3 - ax_6^3).$$

For any $p \in \mathcal{S}$, the equation $\Psi(x_1, \dots, x_6) = 0$ has no solution in $\mathbb{Q}_p$ other than the trivial one, and hence the same is true of the pair of equations

$$\Psi(x_1, \dots, x_6) = \Psi(x_7, \dots, x_{12}) = 0.$$

In addition, the 7-adic condition in the statement of Theorem 1 cannot be removed. Davenport and Lewis [12] observed that when

$$\begin{aligned} \Xi(x_1, \dots, x_5) &= x_1^3 + 2x_2^3 + 6x_3^3 - 4x_4^3, \\ \mathrm{H}(x_1, \dots, x_5) &= \qquad\ \ x_2^3 + 2x_3^3 + 4x_4^3 + x_5^3, \end{aligned}$$

then the pair of equations in 15 variables given by

$$\begin{aligned} \Xi(x_1, \dots, x_5) + 7\Xi(x_6, \dots, x_{10}) + 49\Xi(x_{11}, \dots, x_{15}) &= 0 \\ \mathrm{H}(x_1, \dots, x_5) + 7\mathrm{H}(x_6, \dots, x_{10}) + 49\mathrm{H}(x_{11}, \dots, x_{15}) &= 0 \end{aligned}$$

has no non-trivial solutions in $\mathbb{Q}_7$. In view of these examples, the state of knowledge concerning the local solubility of systems of the type (1.1) may be regarded as having been satisfactorily resolved in all essentials by Davenport and Lewis, and by Cook, at least when $s \geq 13$. Davenport and Lewis [12] showed first that whenever $s \geq 16$, there are non-trivial solutions of (1.1) in any p-adic field. Later, Cook [11] confirmed that such remains true for $13 \leq s \leq 15$ provided only that $p \neq 7$.

Our proof of Theorem 1 uses analytic tools, and in particular employs the circle method. It is a noteworthy feature of our techniques that the method, when it succeeds at all, provides a lower bound for the number of integral solutions of (1.1) in a large box that is essentially best possible. In order to be more precise, when P is a positive number, denote by $\mathcal{N}(P)$ the number of integral solutions $(x_1 \dots, x_s)$ of (1.1) with $|x_j| \leq P$ $(1 \leq j \leq s)$. Then provided that there are solutions of (1.1) in every p-adic field, the principles underlying the Hardy-Littlewood method suggest that an asymptotic formula for $\mathcal{N}(P)$ should hold in which the main term is of size P^{s-6}. We are able to confirm the lower bound $\mathcal{N}(P) \gg P^{s-6}$ implicit in the latter prediction whenever the intersection (1.1) is in general position. This observation is made precise in the following theorem.

Theorem 2. *Let s be a natural number with $s \geq 13$. Suppose that $a_i, b_i \in \mathbb{Z}$ $(1 \leq i \leq s)$ satisfy the condition that for any pair $(c, d) \in \mathbb{Z}^2 \setminus \{(0,0)\}$, at least eight of the numbers $ca_j + db_j$ $(1 \leq j \leq s)$ are non-zero. Then provided that the system* (1.1) *has a non-trivial* 7*-adic solution, one has $\mathcal{N}(P) \gg P^{s-6}$.*

The methods employed by earlier writers, with the exception of Cook [10], were not of sufficient strength to provide a lower bound for $\mathcal{N}(P)$ attaining the order of magnitude presumed to reflect the true state of affairs.

The expectation discussed in the preamble to the statement of Theorem 2 explains the presumed impossibility of a successful application of the circle method to establish analogues of Theorems 1 and 2 with the condition $s \geq 13$ relaxed to the weaker constraint $s \geq 12$. For it is inherent in applications of the circle method to problems involving equations of degree exceeding 2 that error terms arise of size exceeding the square-root of the number of choices for all of the underlying variables. In the context of Theorem 2, the latter error term will exceed a quantity of order $P^{s/2}$, while the anticipated main term in the asymptotic formula for $\mathcal{N}(P)$ is of order P^{s-6}. It is therefore apparent that this latter term cannot be expected to majorize the error term when $s \leq 12$.

The conclusion of Theorem 2 is susceptible to some improvement. The hypotheses can be weakened so as to require that only seven of the numbers $ca_j + db_j$ $(1 \leq j \leq s)$ be non-zero for all pairs $(c, d) \in \mathbb{Z}^2 \setminus \{(0,0)\}$; however, the extra cases would involve us in a lengthy additional discussion within the circle method analysis to follow, and as it stands, Theorem 2 suffices for our immediate purpose at hand. For a refinement of Theorem 2 along these lines, we refer the reader to our forthcoming communication [8].

In the opposite direction, we note that the lower bound recorded in the statement of Theorem 2 is not true without some condition on the coefficients of the type currently imposed. In order to see this, consider the form $\Psi(\mathbf{x})$ defined by

$$\Psi(x_1, x_2, x_3, x_4) = 5x_1^3 + 9x_2^3 + 10x_3^3 + 12x_4^3.$$

Cassels and Guy [9] showed that although the equation $\Psi(\mathbf{x}) = 0$ admits non-trivial solutions in every p-adic field, there are no such solutions in rational integers. Consequently, for any choice of coefficients $\mathbf{b} \in (\mathbb{Z} \setminus \{\mathbf{0}\})^s$, the number of solutions $\mathcal{N}(P)$ associated with the pair of equations

$$\Psi(x_1, x_2, x_3, x_4) = b_1x_1^3 + b_2x_2^3 + \ldots + b_sx_s^3 = 0 \tag{1.2}$$

is equal to the number of integral solutions $(x_5, \ldots, x_s)$ of the single equation $b_5x_5^3 + \ldots + b_sx_s^3 = 0$, with $|x_i| \leq P$ $(5 \leq i \leq s)$. For the system (1.2), therefore, it follows from the methods underlying [17] that $\mathcal{N}(P) \asymp P^{s-7}$ whenever $s \geq 12$. In circumstances in which the system (1.2) possesses non-singular p-adic solutions in every p-adic field, the latter is of smaller order than the prediction $\mathcal{N}(P) \asymp P^{s-6}$ consistent with the conclusion of Theorem 2 that is motivated by a consideration of the product of local densities. Despite the abundance of integral solutions of the system (1.2) for $s \geq 12$, weak approximation also fails. In contrast, with some additional work, our proof of Theorem 2 would extend to establish weak approximation for the system (1.1) without any alteration of the conditions currently imposed. Perhaps weak approximation holds for the system (1.1) with the hypotheses of Theorem 2 relaxed so as to require only that for any $(c,d) \in \mathbb{Z}^2 \setminus \{(0,0)\}$, at least five of the numbers $ca_j + db_j$ $(1 \leq j \leq s)$ are non-zero. However, in order to prove such a conclusion, it seems necessary first to establish that weak approximation holds for diagonal cubic equations in five or more variables. Swinnerton-Dyer [15] has recently obtained such a result subject to the as yet unproven finiteness of the Tate-Shafarevich group for elliptic curves over quadratic fields.

This paper is organised as follows. In the next section, we announce the two mean value estimates that embody the key innovations of this paper; these are recorded in Theorems 3 and 4. Next, in section 3, we introduce a new method for averaging Fourier coefficients over thin sequences, and we apply it to establish Theorem 3. Though motivated by recent work of Wooley [25] and Brüdern, Kawada and Wooley [6], this section contains the most novel material in this paper. In section 4, we derive Theorem 4 as well as some other mean value estimates that all follow from Theorem 3. Then, in section 5, we prepare the stage for a performance of the Hardy-Littlewood method that ultimately establishes Theorem 2. The minor arcs require a rather delicate pruning argument that depends heavily on two innovations for smooth cubic Weyl sums from our recent paper [7]. For more detailed comments on this matter, the reader is directed to section 6, where the pruning is executed, and in particular to the comments introducing section 6. The analysis of the major arcs is standard, and deserves only the abbreviated discussion presented in section 7. In the final section, we derive Theorem 1 from Theorem 2.

Throughout, the letter ε will denote a sufficiently small positive number. We use $\ll$ and $\gg$ to denote Vinogradov's well-known notation, implicit constants depending at most on ε, unless otherwise indicated. In an effort to simplify our analysis, we adopt the convention that whenever ε appears in a statement, then we are implicitly asserting that for each $\varepsilon > 0$ the statement holds for sufficiently large values of the main parameter. Note that the "value" of ε may consequently change from statement to statement, and hence also the dependence of implicit constants on ε. Finally, from time to time we make use of vector notation in order to save space. Thus, for example, we may abbreviate $(c_1, \ldots, c_t)$ to $\mathbf{c}$.

2. A twelfth moment of cubic Weyl sums. In this section we describe the new ingredients employed in our application of the Hardy-Littlewood method to prove Theorem 2. The success of the method depends to a large extent on a new mean value estimate for cubic Weyl sums that we now describe. When P and R are real numbers with $1 \leq R \leq P$, define the set of smooth numbers $\mathcal{A}(P,R)$ by

$$\mathcal{A}(P,R) = \{n \in \mathbb{N} \cap [1,P] : p|n \text{ implies } p \leq R\},$$

where, here and later, the letter p is reserved to denote a prime number. The smooth Weyl sum $h(\alpha) = h(\alpha; P, R)$ central to our arguments is defined by

$$h(\alpha; P, R) = \sum_{x \in \mathcal{A}(P,R)} e(\alpha x^3),$$

where here and hereafter we write $e(z)$ for $e^{2\pi i z}$. An upper bound for the sixth moment of this sum is crucial for the discourse to follow. In order to make our conclusions amenable to possible future progress, we formulate the main estimate explicitly in terms of the sixth moment of $h(\alpha)$. It is therefore convenient to refer to an exponent ξ as *admissible* if, for each positive number ε, there exists a positive number $\eta = \eta(\varepsilon)$ such that, whenever $1 \le R \le P^\eta$, one has the estimate

$$\int_0^1 |h(\alpha; P, R)|^6 \, d\alpha \ll P^{3+\xi+\varepsilon}. \tag{2.1}$$

Lemma 1. *The number* $\xi = (\sqrt{2833} - 43)/41$ *is admissible.*

This is the main result of [22]. Since $(\sqrt{2833} - 43)/41 = 0.2494\ldots$, it follows that there exist admissible exponents ξ with $\xi < 1/4$, a fact of importance to us later. The first admissible exponent smaller than $1/4$ was obtained by Wooley [21].

Next, when $a, b, c, d \in \mathbb{Z}$ and $\mathcal{B}$ is a finite set of integers, we define the integral

$$I(a, b, c, d) = \int_0^1 \int_0^1 |h(a\alpha)h(b\beta)|^5 \Big| \sum_{z \in \mathcal{B}} e\big((c\alpha + d\beta)z^3\big)\Big|^2 \, d\alpha \, d\beta. \tag{2.2}$$

We may now announce our central auxiliary mean value estimate, which we prove in section 3.

Theorem 3. *Suppose that* a, b, c, d *are non-zero integers, and that* $\mathcal{B} \subseteq [1, P] \cap \mathbb{Z}$. *Then for each admissible exponent* ξ, *and for each positive number* ε, *there exists a positive number* $\eta = \eta(\varepsilon)$ *such that, whenever* $1 \le R \le P^\eta$, *one has*

$$I(a, b, c, d) \ll P^{6+\xi+\varepsilon}.$$

If one takes $\mathcal{B} = \mathcal{A}(P, R)$, then the conclusion of Theorem 3 yields the estimate

$$\int_0^1 \int_0^1 |h(a\alpha)^5 h(b\beta)^5 h(c\alpha + d\beta)^2| \, d\alpha \, d\beta \ll P^{6+\xi+\varepsilon}. \tag{2.3}$$

While this bound suffices for the applications discussed in this paper, the more general conclusion recorded in Theorem 3 is required in our forthcoming article [8]. We note that previous writers would apply Hölder's inequality and suitable changes of variable so as to bound the left hand side of (2.3) in terms of factorisable double integrals of the shape

$$\int_0^1 \int_0^1 |h(A\alpha)h(B\beta)|^6 \, d\alpha \, d\beta, \tag{2.4}$$

with suitable fixed integers A and B satisfying $AB \neq 0$. The latter integral may be estimated via the inequality (2.1), and thereby workers hitherto would derive an upper bound of the shape (2.3), but with the exponent $6 + 2\xi + \varepsilon$ in place of $6 + \xi + \varepsilon$. Underpinning these earlier strategies are mean values involving two linearly independent linear forms in α and β, these being reducible to the shape (2.4). In contrast, our approach in this paper makes crucial use of the presence within the mean value (2.3) of three pairwise linearly independent linear forms in α and β, and we save a factor of P^ξ by exploiting the extra structure inherent in such mean values. It is worth noting that the existence of an upper bound for the mean value (2.4) of order $P^{6+2\xi+\varepsilon}$ is essentially equivalent to the validity of the estimate (2.1), and thus the strategy underlying the proof of Theorem 3 is inherently superior to that applied by previous authors whenever the sharpest available admissible exponent ξ is non-zero.

As another corollary of Theorem 3, we derive a more symmetric twelfth moment estimate in section 4 below.

Theorem 4. *Suppose that c_i, d_i $(1 \le i \le 3)$ are integers satisfying the condition*

$$(c_1d_2 - c_2d_1)(c_1d_3 - c_3d_1)(c_2d_3 - c_3d_2) \neq 0. \tag{2.5}$$

Write $\Lambda_j = c_j\alpha + d_j\beta$ $(1 \le j \le s)$. Then for each admissible exponent ξ, and for each positive number ε, there exists a positive number $\eta = \eta(\varepsilon)$ such that, whenever $1 \le R \le P^\eta$, one has the estimates

$$\int_0^1\int_0^1 |h(\Lambda_1)^5h(\Lambda_2)^5h(\Lambda_3)^2|\, d\alpha\, d\beta \ll P^{6+\xi+\varepsilon} \tag{2.6}$$

and

$$\int_0^1\int_0^1 |h(\Lambda_1)h(\Lambda_2)h(\Lambda_3)|^4\, d\alpha\, d\beta \ll P^{6+\xi+\varepsilon}. \tag{2.7}$$

Note that the integral estimated in (2.7) has a natural interpretation as the number of solutions of a pair of diophantine equations, an advantageous feature absent from both (2.3) and (2.6). We remark also that conclusions analogous to those recorded in Theorems 3 and 4 may be derived with the cubic exponential sums replaced by sums of higher degree. Indeed, both the conclusions and their proofs are essentially identical with those presented in this paper, save that the admissible exponent ξ herein is replaced by one depending on the degree in question.

3. Averaging Fourier coefficients over thin sequences. Our objective in this section is the proof of Theorem 3. We assume throughout that the hypotheses of the statement of Theorem 3 are satisfied. Thus, in particular, we may suppose that ξ is admissible, and that $\eta = \eta(\varepsilon)$ is a positive number sufficiently small that the estimate (2.1) holds. When $n \in \mathbb{Z}$, we let $r(n)$ denote the number of representations of n in the form $n = x^3 - y^3$, with $x, y \in \mathcal{B}$. It follows that

$$\Big|\sum_{z\in\mathcal{B}} e(\gamma z^3)\Big|^2 = \sum_{|n|\le P^3} r(n)e(-\gamma n). \tag{3.1}$$

We apply this formula to achieve a simple preliminary transformation of the integral $I(a,b,c,d)$ defined in (2.2). In this context, when $l \in \mathbb{Z}$ we write

$$\psi_l(m) = \int_0^1 |h(l\alpha)|^5 e(-\alpha m)\,d\alpha. \tag{3.2}$$

Given $\mathcal{B} \subseteq [1,P] \cap \mathbb{Z}$, the application of (3.1) within (2.2) leads to the relation

$$\begin{aligned} I(a,b,c,d) &= \sum_{|n|\le P^3} r(n) \int_0^1 \int_0^1 |h(a\alpha)|^5 |h(b\beta)|^5 e(-cn\alpha)e(-dn\beta)\,d\alpha\,d\beta \\ &= \sum_{|n|\le P^3} r(n)\psi_a(cn)\psi_b(dn). \end{aligned}$$

Observe from (3.2) that $\psi_l(m)$ is real for any pair of integers l and m. Then by Cauchy's inequality, we derive the basic estimate

$$I(a,b,c,d) \ll \Bigg(\sum_{|n|\le P^3} r(n)\psi_a(cn)^2 \Bigg)^{1/2} \Bigg(\sum_{|n|\le P^3} r(n)\psi_b(dn)^2 \Bigg)^{1/2}. \tag{3.3}$$

Further progress now depends on a new method for counting integers in thin sequences for which certain arithmetically defined Fourier coefficients are abnormally large. Recent work of Wooley [25] provides a framework for providing good estimates for the number of integers having unusually many representations as the sum of a fixed number of cubes. In a different direction, the discussion in Brüdern, Kawada and Wooley [6] supplies a strategy for bounding similar exceptional sets over thin sequences. Motivated by such arguments, we study the Fourier coefficients $\psi_l(km)$ for fixed integers l and k, and in Lemma 2 below we estimate the number of occurrences of large values of $|\psi_l(kn)|$ as n varies over the set $\mathcal{Z} = \{n \in \mathbb{Z} : r(n) > 0\}$. This information is then converted, in Lemma 3, into a mean square bound for $\psi_l(kn)$ averaged over $\mathcal{Z}$. Suitably positioned to bound the sums on the right hand side of (3.3), the proof of Theorem 3 is swiftly completed.

Before advancing to establish Lemma 2, we require some notation. When l and k are fixed integers and T is a non-negative real number, we define the set $\mathcal{Z}(T) = \mathcal{Z}_{l,k}(T)$ by

$$\mathcal{Z}_{l,k}(T) = \{n \in \mathcal{Z} : |\psi_l(kn)| > T\}.$$

For the remainder of this section we assume that our basic parameter P is a large positive number, and that l and k are fixed non-zero integers.

Lemma 2. *Whenever δ is a positive number and $T \ge P^{2+\xi/2+\delta}$, one has the upper bound card$(\mathcal{Z}(T)) \ll P^{6+\xi+\varepsilon}T^{-2}$.*

Proof. We define the coefficient σ_m for each integer m by means of the relation $\psi_l(m) = \sigma_m|\psi_l(m)|$ when $\psi_l(m) \ne 0$, and otherwise by putting $\sigma_m = 0$. Since $\mathcal{Z} \subseteq [-P^3, P^3]$, we can define the finite exponential sum

$$K_T(\alpha) = \sum_{n\in\mathcal{Z}(T)} \sigma_{kn} e(-kn\alpha).$$

In view of (3.2), it follows that

$$\sum_{n\in\mathcal{Z}(T)} |\psi_l(kn)| = \int_0^1 |h(l\alpha)|^5 K_T(\alpha)d\alpha. \tag{3.4}$$

At this point, in the interest of brevity, we write $Z_T = \mathrm{card}(\mathcal{Z}(T))$. Then the left hand side of (3.4) must exceed TZ_T, whence Schwarz's inequality yields the bound

$$TZ_T \le \left(\int_0^1 |h(l\alpha)|^6 d\alpha\right)^{1/2} \left(\int_0^1 |h(l\alpha)^4 K_T(\alpha)^2| d\alpha\right)^{1/2}. \tag{3.5}$$

By (2.1) and a transparent change of variable, the first integral on the right hand side of (3.5) is $O(P^{3+\xi+\varepsilon})$. In order to estimate the second integral, one first applies Weyl's differencing lemma to $|h(l\alpha)|^4$ (see Lemma 2.3 of [19]), and then interprets the resulting expression in terms of the underlying diophantine equation. Thus, one obtains

$$\int_0^1 |h(l\alpha)^4 K_T(\alpha)^2| d\alpha \ll P^\varepsilon (P^3 Z_T + P Z_T^2). \tag{3.6}$$

For full details of this estimation, we refer the reader to Lemma 2.1 of Wooley [24], where a proof is described in the special case $l = 1$ that readily extends to the present situation. As an alternative, we direct the reader to the method of proof of Lemma 5.1 of [23]. Collecting together (3.5) and (3.6), we conclude that

$$TZ_T \ll P^{3/2+\xi/2+\varepsilon}(P^3 Z_T + P Z_T^2)^{1/2} = P^{3+\xi/2+\varepsilon} Z_T^{1/2} + P^{2+\xi/2+\varepsilon} Z_T.$$

The proof of the lemma is completed by recalling our assumption that $T > P^{2+\xi/2+\delta}$, where δ is a positive number that we may suppose to exceed 2ε.

Lemma 3. *One has* $\sum_{n\in\mathcal{Z}} \psi_l(kn)^2 \ll P^{6+\xi+\varepsilon}$.

Proof. Our discussion is facilitated by a division of the set $\mathcal{Z}$ into various subsets. To this end, we fix a positive number δ and define

$$\mathcal{Y}_0 = \{n \in \mathcal{Z} : |\psi_l(kn)| \le P^{2+\xi/2+\delta}\}. \tag{3.7}$$

Also, when $T \ge 1$, we put $\mathcal{Y}(T) = \{n \in \mathcal{Z} : T < |\psi_l(kn)| \le 2T\}$. On noting the trivial upper bound $\mathrm{card}(\mathcal{Z}) \le P^2$, it is apparent from (3.7) that

$$\sum_{n\in\mathcal{Y}_0} \psi_l(kn)^2 \le P^2 (P^{2+\xi/2+\delta})^2 \ll P^{6+\xi+2\delta}. \tag{3.8}$$

The bound $|\psi_l(kn)| \le P^5$, on the other hand, valid uniformly for $n \in \mathbb{Z}$, follows from (3.2) via the triangle inequality. A familiar argument involving a dyadic dissection therefore establishes that for some number T with $P^{2+\xi/2+\delta} \le T \le P^5$, one has

$$\sum_{n\in\mathcal{Z}} \psi_l(kn)^2 \ll \sum_{n\in\mathcal{Y}_0} \psi_l(kn)^2 + (\log P) \sum_{n\in\mathcal{Y}(T)} \psi_l(kn)^2. \tag{3.9}$$

But $\mathcal{Y}(T) \subseteq \mathcal{Z}(T)$, and so it follows from Lemma 2 that

$$\sum_{n \in \mathcal{Y}(T)} \psi_l(kn)^2 \leq (2T)^2 \text{card}(\mathcal{Z}(T)) \ll P^{6+\xi+\varepsilon}. \tag{3.10}$$

The conclusion of Lemma 3 is obtained by substituting (3.8) and (3.10) into (3.9), and then taking $\delta = \varepsilon/2$.

Lemma 4. *One has* $\sum_{|n| \leq P^3} r(n)\psi_l(kn)^2 \ll P^{6+\xi+\varepsilon}$.

Proof. We begin by noting that a simple divisor argument shows that whenever m is a non-zero integer, then $r(m) = O(P^\varepsilon)$. Since also $r(0) \leq P$, we find that

$$\sum_{|n| \leq P^3} r(n)\psi_l(kn)^2 \ll P\psi_l(0)^2 + P^\varepsilon \sum_{n \in \mathcal{Z}} \psi_l(kn)^2. \tag{3.11}$$

On recalling (3.2), moreover, it follows from a change of variable in combination with Schwarz's inequality that

$$\psi_l(0) = \int_0^1 |h(l\alpha)|^5 d\alpha \leq \left(\int_0^1 |h(\alpha)|^4 d\alpha \right)^{1/2} \left(\int_0^1 |h(\alpha)|^6 d\alpha \right)^{1/2}. \tag{3.12}$$

The first integral on the right hand side of (3.12) may be estimated by means of Hua's Lemma (see [19, Lemma 2.5]), and the second via (2.1). Thus we find that

$$\psi_l(0)^2 \ll (P^{2+\varepsilon})(P^{3+\xi+\varepsilon}) = P^{5+\xi+2\varepsilon}.$$

The proof of the lemma is completed by substituting the latter bound, together with the estimate provided by Lemma 3, into the relation (3.11).

In order to establish Theorem 3, we have merely to apply Lemma 4 with (l, k) equal to (a, c) and (b, d) respectively, and then make use of the inequality (3.3).

4. Some mean value estimates. At this point it is convenient to explore some consequences of Theorem 3 that are relevant for our later proceedings. We suppose throughout this section that ξ is admissible, and that $\eta = \eta(\varepsilon)$ is a positive number sufficiently small that the estimate (2.1) holds. We begin by deriving Theorem 4, and here we make use of the notation introduced in the statement of this theorem presented in section 2.

The proof of Theorem 4. When k is an integer with $1 \leq k \leq 3$, let i and j be the integers for which $\{i, j, k\} = \{1, 2, 3\}$, and write

$$J_k = \int_0^1 \int_0^1 |h(\Lambda_i)^5 h(\Lambda_j)^5 h(\Lambda_k)^2| \, d\alpha \, d\beta. \tag{4.1}$$

Then it follows from Hölder's inequality that

$$\int_0^1\int_0^1 |h(\Lambda_1)h(\Lambda_2)h(\Lambda_3)|^4 d\alpha\, d\beta \le (J_1J_2J_3)^{1/3}. \tag{4.2}$$

The conclusion of Theorem 4 is immediate from the estimate $J_k = O(P^{6+\xi+\varepsilon})$ $(1\le k\le 3)$, which we now seek to establish.

By way of example we estimate J_3. Corresponding estimates for J_1 and J_2 follow by symmetrical arguments. We begin by observing that the hypotheses of Theorem 4 ensure that any two of the linear forms Λ_1, Λ_2 and Λ_3 are linearly independent, whence there are non-zero integers A, B and C, depending at most on $\mathbf{c}$ and $\mathbf{d}$, for which $C\Lambda_3 = A\Lambda_1 + B\Lambda_2$. Making use of the periodicity (with period 1) of the integrand in (4.1), and changing variables, one therefore finds that

$$\begin{aligned} J_3 &= C^{-2}\int_0^C\int_0^C |h(\Lambda_1)^5h(\Lambda_2)^5h(\Lambda_3)^2|\, d\alpha\, d\beta \\ &= \int_0^1\int_0^1 |h(C\Lambda_1)^5h(C\Lambda_2)^5h(A\Lambda_1+B\Lambda_2)^2|\, d\alpha\, d\beta. \end{aligned} \tag{4.3}$$

Now change the variables of integration from (α,β) to (Λ_1,Λ_2), and observe that the resulting range of integration becomes a parallelogram contained in a square with sides of integral length parallel to the coordinate axes of the (Λ_1,Λ_2)-plane. Plainly, moreover, the dimensions of this square depend at most on $\mathbf{c}$ and $\mathbf{d}$. Making use again of the periodicity (with period 1) of the integrand, we thus obtain the estimate

$$J_3 \ll \int_0^1\int_0^1 |h(C\Lambda_1)^5h(C\Lambda_2)^5h(A\Lambda_1+B\Lambda_2)^2|\, d\Lambda_1\, d\Lambda_2.$$

The upper bound $J_3 = O(P^{6+\xi+\varepsilon})$ now follows from the consequence (2.3) of Theorem 3, and on making use of the corresponding symmetrical bounds for J_1 and J_2, the conclusion of Theorem 4 is immediate from (4.2).

In preparation for the next lemma, we record an elementary estimate of utility in the arguments to follow that involve some level of combinatorial complexity.

Lemma 5. *Let k and N be natural numbers, and suppose that $\mathfrak{B}\subseteq\mathbb{C}^k$ is measurable. Let $u_i(\mathbf{z})$ $(0\le i\le N)$ be complex-valued functions of $\mathfrak{B}$. Then whenever the functions $|u_0(\mathbf{z})u_j(\mathbf{z})^N|$ $(1\le j\le N)$ are integrable on $\mathfrak{B}$, one has the upper bound*

$$\int_{\mathfrak{B}} |u_0(\mathbf{z})u_1(\mathbf{z})\dots u_N(\mathbf{z})|\, d\mathbf{z} \le N \max_{1\le j\le N}\int_{\mathfrak{B}} |u_0(\mathbf{z})u_j(\mathbf{z})^N|d\mathbf{z}.$$

Proof. The desired conlcusion is immediate from the inequality $|\zeta_1\zeta_2\dots\zeta_N| \le |\zeta_1|^N + |\zeta_2|^N + \dots + |\zeta_N|^N$ that is valid for any complex numbers ζ_i $(1\le i\le N)$.

The next lemma contains (2.3) and Theorem 4 as special cases, and yet has a shape sufficiently general that it may be easily applied in what follows. In order to describe the conclusion of this lemma, we consider integers c_j and d_j with $(c_j, d_j) \neq (0,0)$ $(1 \leq j \leq 12)$. To each pair (c_j, d_j) we associate the linear form $\Lambda_j = c_j\alpha + d_j\beta$. We describe two such forms Λ_i and Λ_j as *equivalent* when there exists a non-zero rational number λ with $\Lambda_i = \lambda\Lambda_j$. This notion plainly defines an equivalence relation on the set $\{\Lambda_1, \ldots, \Lambda_{12}\}$, and we refer to the number of elements in the equivalence class containing the form Λ_j as its *multiplicity.* Finally, in order to promote concision, for each index l we abbreviate $|h(\Lambda_l)|$ simply to h_l.

Lemma 6. *In the setting described in the preamble to this lemma, suppose that the multiplicities of the linear forms* $\Lambda_1, \ldots, \Lambda_{12}$ *are at most* 5*. Then*

$$\int_0^1 \int_0^1 h_1 h_2 \ldots h_{12}\, d\alpha\, d\beta \ll P^{6+\xi+\varepsilon}.$$

Proof. Consider the situation in which the number of equivalence classes amongst $\Lambda_1, \ldots, \Lambda_{12}$ is t. By relabelling indices if necessary, we may suppose that representatives of these equivalence classes are $\Lambda_1, \ldots, \Lambda_t$. For each index i, let r_i denote the number of linear forms amongst $\Lambda_1, \ldots, \Lambda_{12}$ equivalent to Λ_i. Then in view of the hypotheses of the lemma, we may relabel indices so as to ensure that

$$1 \leq r_t \leq r_{t-1} \leq \ldots \leq r_1 \leq 5 \quad \text{and} \quad r_1 + r_2 + \cdots + r_t = 12. \tag{4.4}$$

Next, for a given index i with $1 \leq i \leq t$, consider the linear forms Λ_{l_j} $(1 \leq j \leq r_i)$ equivalent to Λ_i. Apply Lemma 5 with $N = r_i$, with h_{l_j} in place of u_j $(1 \leq j \leq N)$, and with u_0 replaced by the product of those h_l with Λ_l not equivalent to Λ_i. Then it is apparent that there is no loss of generality in supposing that $\Lambda_{l_j} = \Lambda_i$ $(1 \leq j \leq r_i)$. By repeating this argument for successive equivalence classes, moreover, we find that

$$\int_0^1 \int_0^1 h_1 \ldots h_{12}\, d\alpha\, d\beta \ll \int_0^1 \int_0^1 h_1^{r_1} \ldots h_t^{r_t}\, d\alpha\, d\beta. \tag{4.5}$$

A further simplification neatly sidesteps combinatorial complications. Let ν be a non-negative integer, and suppose that $r_{t-1} = r_t + \nu < 5$. Then we may apply Lemma 5 with $N = \nu + 2$, with h_{t-1} in place of u_i $(1 \leq i \leq \nu + 1)$ and h_t in place of u_N, and with u_0 set equal to

$$h_1^{r_1} h_2^{r_2} \ldots h_{t-2}^{r_{t-2}} h_{t-1}^{r_{t-1}-\nu-1} h_t^{r_t-1}.$$

Here, and in what follows, we interpret the vanishing of any exponent as indicating that the associated exponential sum is deleted from the product. In this way we obtain an upper bound of the shape (4.5) in which the exponents r_{t-1} and $r_t = r_{t-1} - \nu$ are replaced by $r_{t-1} + 1$ and $r_t - 1$, respectively, or else by $r_{t-1} - \nu - 1$ and $r_t + \nu + 1$. By relabelling if necessary, we derive an upper bound of the shape (4.5), subject to the constraints (4.4), wherein either the parameter r_t is reduced, or else the parameter t is reduced. By repeating this process, therefore, we ultimately arrive at a situation in which $r_{t-1} = 5$, and then the constraints (4.4) imply that necessarily $(r_1, r_2, \dots, r_t) = (5, 5, 2)$. The conclusion of the lemma is now immediate from (4.5) on making use of the estimate (2.6) of Theorem 4.

As is often the case with applications of the circle method, it is desirable to have available a sharp upper bound bought with additional generating functions. We begin with an auxiliary lemma analogous to Theorem 4. In this context we take $\mathfrak{m}$ to be the set of real numbers $\alpha \in [0,1)$ such that, whenever $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $(a,q) = 1$ and $|q\alpha - a| \le P^{-9/4}$, then one has $q > P^{3/4}$. We then put $\mathfrak{M} = [0,1) \setminus \mathfrak{m}$.

Lemma 7. *Suppose that c_i, d_i $(1 \le i \le 3)$ are integers satisfying the condition* (2.5)*. Then, in the notation employed in the statement of Theorem* 4*, one has*

$$\int_0^1 \int_0^1 h_1^6 h_2^6 h_3^2 \, d\alpha \, d\beta \ll P^8.$$

Proof. We observe that the argument leading from (4.1) to (4.3) reveals first that there are non-zero integers A, B and C for which $C\Lambda_3 = A\Lambda_1 + B\Lambda_2$, and then via a change of variables that

$$\int_0^1 \int_0^1 h_1^6 h_2^6 h_3^2 \, d\alpha \, d\beta \ll \mathcal{I}_1(A, B, C), \tag{4.6}$$

where we write

$$\mathcal{I}_1(A, B, C) = \int_0^1 \int_0^1 |h(C\alpha)^6 h(C\beta)^6 h(A\alpha + B\beta)^2| \, d\alpha \, d\beta. \tag{4.7}$$

By orthogonality, the mean value $\mathcal{I}_1(A, B, C)$ is bounded above by the number of integral solutions of the diophantine system

$$A^{-1} \sum_{i=1}^{3} (x_{2i-1}^3 - x_{2i}^3) = B^{-1} \sum_{i=1}^{3} (y_{2i-1}^3 - y_{2i}^3) = C^{-1}(z_1^3 - z_2^3),$$

with $1 \le x_1, y_1 \le P$ and $x_j, y_j, z_l \in \mathcal{A}(P, R)$ $(2 \le j \le 6,\ 1 \le l \le 2)$. We now introduce the classical Weyl sum

$$f(\alpha) = \sum_{1 \le x \le P} e(\alpha x^3),$$

and define the mean value $\mathcal{I}_2(\mathfrak{B}) = \mathcal{I}_2(\mathfrak{B}; A, B, C)$, for a measurable set $\mathfrak{B}$, by

$$\mathcal{I}_2(\mathfrak{B}) = \iint_{\mathfrak{B}} |f(C\alpha) f(C\beta) h(C\alpha)^5 h(C\beta)^5 h(A\alpha + B\beta)^2| \, d\alpha \, d\beta. \tag{4.8}$$

Then on applying orthogonality in combination with the triangle inequality, we may conclude that

$$\mathcal{I}_1 \le \mathcal{I}_2([0,1)^2). \tag{4.9}$$

We estimate the integral on the right hand side of (4.8) by a simple version of the circle method. By an enhanced version of Weyl's inequality (see [17, Lemma 1]), one readily confirms that whenever a is a fixed non-zero integer, then

$$\sup_{\theta\in\mathfrak{m}} |f(a\theta)| \ll P^{3/4+\varepsilon}. \tag{4.10}$$

In view of the trivial upper bound $|f(\theta)| \le P$, one deduces that when $(\alpha,\beta) \in [0,1)^2$ and the upper bound $|f(C\alpha)f(C\beta)| \ll P^{7/4+\varepsilon}$ fails to hold, then necessarily $(\alpha,\beta) \in \mathfrak{M}^2$. Consequently, it follows from (4.8) and (4.9) that

$$\mathcal{I}_1 \ll P^{7/4+\varepsilon} \int_0^1\int_0^1 |h(C\alpha)^5 h(C\beta)^5 h(A\alpha+B\beta)^2|\, d\alpha\, d\beta + \mathcal{I}_2(\mathfrak{M}^2). \tag{4.11}$$

On recalling (4.7) and applying Hölder's inequality to (4.8), one finds that

$$\mathcal{I}_2(\mathfrak{M}^2) \le \mathcal{I}_1^{5/6} \Big(\iint_{\mathfrak{M}^2} |f(C\alpha)^6 f(C\beta)^6 h(A\alpha+B\beta)^2|\, d\alpha\, d\beta \Big)^{1/6}.$$

A standard application of the Hardy-Littlewood method (see Chapter 4 of [19]), moreover, readily confirms that whenever a is a fixed non-zero integer, one has

$$\int_{\mathfrak{M}} |f(a\theta)|^6\, d\theta \ll P^3.$$

Thus, on making use of the trivial bound $|h(A\alpha+B\beta)| \le P$, we see that

$$\mathcal{I}_2(\mathfrak{M}^2) \ll \mathcal{I}_1^{5/6}(P^8)^{1/6}.$$

On substituting the latter relation into (4.11) and recalling the estimate (2.3), we deduce that for a suitably small positive number δ, one has

$$\mathcal{I}_1 \ll P^{8-\delta} + P^{4/3}\mathcal{I}_1^{5/6},$$

whence $\mathcal{I}_1 \ll P^8$. The conclusion of the lemma is now immediate fom (4.6).

With greater effort one may establish an asymptotic formula for the mean value recorded in the statement of Lemma 7, thereby confirming that the upper bound therein is of the correct order of magnitude. Were our estimate to be weaker by a factor of P^ε, our subsequent deliberations would be greatly complicated.

5. Preparing the stage for Hardy and Littlewood. We are now equipped with auxiliary mean value estimates sufficient for our intended task, and so we return to our main concern and count integral solutions of the system (1.1) via the Hardy-Littlewood method. We suppose that the hypotheses of the statement of Theorem 2 are satisfied, so that, in particular, one has $s \ge 13$. With the pairs $(a_j,b_j) \in \mathbb{Z}^2 \setminus \{(0,0)\}$ $(1 \le j \le s)$, we associate both the linear forms

$$\Lambda_j = a_j\alpha + b_j\beta \quad (1 \le j \le s), \tag{5.1}$$

and the two linear forms $L_1(\boldsymbol{\theta})$ and $L_2(\boldsymbol{\theta})$ defined for $\boldsymbol{\theta} \in \mathbb{R}^s$ by

$$L_1(\boldsymbol{\theta}) = \sum_{j=1}^{s} a_j\theta_j \quad \text{and} \quad L_2(\boldsymbol{\theta}) = \sum_{j=1}^{s} b_j\theta_j. \tag{5.2}$$

Recall the notions of equivalence and multiplicity of linear forms from the preamble to Lemma 6, and extend these conventions in the natural way so as to apply to the set $\{\Lambda_1, \ldots, \Lambda_s\}$. By the hypotheses of the statement of Theorem 2, one finds that for any pair $(c, d) \in \mathbb{Z}^2 \setminus \{(0, 0)\}$, the linear form $cL_1(\boldsymbol{\theta}) + dL_2(\boldsymbol{\theta})$ necessarily posesses at least eight non-zero coefficients. By choosing an appropriate subset $\mathcal{S}$ of $\{1, \ldots, s\}$ with $\mathrm{card}(\mathcal{S}) = 13$, we may therefore ensure that at most five of the forms Λ_j with $j \in \mathcal{S}$ belong to the same equivalence class. Suppose that these 13 forms fall into t equivalence classes, and that the multiplicities of the representatives of these classes are $r_1, \ldots, r_t$. In view of our earlier observations, there is no loss of generality in supposing that $5 \geq r_1 \geq r_2 \geq \ldots \geq r_t$ and $r_1 + \cdots + r_t = 13$, and hence, in addition, that $t \geq 3$. With the aim of simplifying our notation, we now relabel variables in the system (1.1), and likewise in (5.1) and (5.2), so that the set $\mathcal{S}$ becomes $\{1, 2, \ldots, 13\}$, and so that Λ_1 becomes a linear form in the first equivalence class counted by r_1, and Λ_2 becomes a form in the second equivalence class counted by r_2.

Next, on taking suitable integral linear combinations of the equations (1.1), we may suppose without loss that

$$b_1 = a_2 = 0. \tag{5.3}$$

Since we may suppose that $a_1b_2 \neq 0$, it is now apparent that the simultaneous equations

$$L_1(\boldsymbol{\theta}) = L_2(\boldsymbol{\theta}) = 0 \tag{5.4}$$

possess a solution $\boldsymbol{\theta}$ with $\theta_j \neq 0$ $(1 \leq j \leq s)$. We next apply the substitution $x_j \to -x_j$ for those indices j with $1 \leq j \leq s$ for which $\theta_j < 0$. Neither the solubility of the system (1.1), nor the corresponding counting function $\mathcal{N}(P)$, are affected by this manoeuvre, and yet the transformed linear system associated with (5.4) has a solution $\boldsymbol{\theta}$ with $\theta_j > 0$ $(1 \leq j \leq s)$. The homogeneity of the system (5.4) ensures, moreover, that a solution of the latter type may be chosen with $\boldsymbol{\theta} \in (0, 1)^s$. We now fix this solution $\boldsymbol{\theta}$, and fix also ε to be a sufficiently small positive number, and η to be a positive number sufficiently small in the context of Theorems 3 and 4 and the associated auxiliary mean value estimates, and so small that one has also $\eta < \theta_j < 1$ $(1 \leq j \leq s)$. In this way, we may suppose that the solution $\boldsymbol{\theta}$ of the linear system (5.4) satisfies $\boldsymbol{\theta} \in (\eta, 1)^s$.

We are at last prepared to describe our strategy for proving Theorem 2. We take P to be a positive number sufficiently large in terms of ε, η, $\mathbf{a}$, $\mathbf{b}$ and $\boldsymbol{\theta}$, and we put $R = P^\eta$. On defining the exponential sum

$$g(\alpha) = \sum_{\eta P < x \leq P} e(\alpha x^3)$$

and the generating functions

$$H_0(\alpha,\beta)=\prod_{j=2}^{13}h(\Lambda_j)\quad\text{and}\quad H(\alpha,\beta)=\prod_{j=2}^{s}h(\Lambda_j),\tag{5.5}$$

it follows from orthogonality that

$$\mathcal{N}(P)\ge\int_0^1\int_0^1 g(\Lambda_1)H(\alpha,\beta)\,d\alpha\,d\beta.\tag{5.6}$$

We analyse the double integral in (5.6) by means of the Hardy-Littlewood method. In this context, we put

$$Q=(\log P)^{1/100},\tag{5.7}$$

and when $a,b\in\mathbb{Z}$ and $q\in\mathbb{N}$, we define the boxes

$$\mathfrak{N}(q,a,b)=\{(\alpha,\beta)\in[0,1)^2\,:\,|\alpha-a/q|\le QP^{-3}\text{ and }|\beta-b/q|\le QP^{-3}\}.$$

Our Hardy-Littlewood dissection is then defined by taking the set $\mathfrak{N}$ of *major arcs* to be the union of the boxes $\mathfrak{N}(q,a,b)$ with $0\le a,b\le q\le Q$ subject to $(a,b,q)=1$, and the *minor arcs* $\mathfrak{n}$ to be the complementary set $[0,1)^2\setminus\mathfrak{N}$.

The contribution to the integral in (5.6) arising from the major arcs $\mathfrak{N}$ satisfies the asymptotic lower bound

$$\iint_{\mathfrak{N}} g(\Lambda_1)H(\alpha,\beta)\,d\alpha\,d\beta\gg P^{s-6},\tag{5.8}$$

a fact whose confirmation is the sole objective of section 7 below. The corresponding contribution of the minor arcs $\mathfrak{n}$ is asymptotically smaller. Indeed, in section 6 we show that

$$\iint_{\mathfrak{n}} g(\Lambda_1)H(\alpha,\beta)\,d\alpha\,d\beta\ll P^{s-6}(\log P)^{-1/140000}.\tag{5.9}$$

The desired conclusion $\mathcal{N}(P)\gg P^{s-6}$ is immediate from (5.8) and (5.9) on recalling that $[0,1)^2$ is the disjoint union of $\mathfrak{N}$ and $\mathfrak{n}$.

6. Pruning to the root. Our goal in this section is the proof of the estimate (5.9). On recalling the definitions (5.5) and making use of the trivial bound $|h(\gamma)|\le P$, it is apparent that the desired estimate follows directly from the following lemma, the proof of which will occupy us for the remainder of this section.

Lemma 8. *Under the hypotheses prevailing in the discourse of section* 5, *one has*

$$\iint_{\mathfrak{n}} |g(\Lambda_1)H_0(\alpha,\beta)|\,d\alpha\,d\beta\ll P^{7}(\log P)^{-1/140000}.\tag{6.1}$$

The proof of Lemma 8 involves an unconventional pruning exercise. One gets started rather easily. Recall the major and minor arcs $\mathfrak{M}$ and $\mathfrak{m}$ introduced in the preamble to Lemma 7, and consider the auxiliary sets

$$\mathfrak{e} = \{(\alpha, \beta) \in \mathfrak{n} : \alpha \in \mathfrak{m}\} \quad \text{and} \quad \mathfrak{E} = \{(\alpha, \beta) \in \mathfrak{n} : \alpha \in \mathfrak{M}\}. \tag{6.2}$$

Then on recalling that $\Lambda_1 = a_1\alpha$, one finds via two applications of (4.10) that

$$\sup_{(\alpha,\beta)\in\mathfrak{e}} |g(\Lambda_1)| = \sup_{\alpha\in\mathfrak{m}} |g(a_1\alpha)| \ll P^{3/4+\varepsilon}.$$

But in view of the definition (5.5), the mean value of $H_0(\alpha, \beta)$ may be estimated by means of Lemma 6. Thus we deduce that

$$\iint_{\mathfrak{e}} |g(\Lambda_1)H_0(\alpha, \beta)|\, d\alpha\, d\beta \ll P^{3/4+\varepsilon} \int_0^1\int_0^1 |H_0(\alpha, \beta)|\, d\alpha\, d\beta \ll P^{27/4+\xi+\varepsilon}. \tag{6.3}$$

The treatment of the complementary set $\mathfrak{E}$ is much harder. Although one already has the potentially powerful information that $\alpha \in \mathfrak{M}$, there is presently no such control available on β. Furthermore, there is only one classical Weyl sum within the product of generating functions on which one may hope to exercise useful control. Nonetheless, we are able to set up machinery with which to prune straight down to the set of narrow arcs $\mathfrak{N}$ by using two different devices from our recent work [7] on cubic smooth Weyl sums. We are very fortunate to be able to borrow from this work, for we have not been successful in constructing an argument of sufficient strength along more conventional lines. Appropriate modifications of the aforementioned devices from [7] are embodied in the following two lemmata.

Lemma 9. *Let A be a fixed non-zero rational number, and let δ be a fixed positive number. Then one has*

$$\sup_{\lambda\in\mathbb{R}} \int_{\mathfrak{M}} |g(a_1\theta)|^{2+\delta} |h(A\theta + \lambda)|^2\, d\theta \ll P^{1+\delta}.$$

It is noteworthy, and important in our later discussion, that the bound here has the expected order of magnitude, uninflated by factors of P^ε. The next lemma shares this feature.

Lemma 10. *(i) One has*

$$\int_0^1 |h(\alpha)|^{77/10}\, d\alpha \ll P^{47/10}.$$

(ii) When Λ_i and Λ_j are inequivalent, one has

$$\iint_{\mathfrak{n}} h_i^8 h_j^8\, d\alpha\, d\beta \ll P^{10} Q^{-3/100}.$$

We postpone the proof of these two lemmata to the end of this section, initiating at once the estimation of the contribution of the set $\mathfrak{E}$ within the mean value on the left hand side of (6.1). Suppose that the number of equivalence classes amongst $\Lambda_2, \ldots, \Lambda_{13}$ is T. By relabelling variables if necessary, we may suppose that representatives of these equivalence classes are $\widetilde{\Lambda}_1, \widetilde{\Lambda}_2, \ldots, \widetilde{\Lambda}_T$. For each index i, let R_i denote the number of linear forms amongst $\Lambda_2, \ldots, \Lambda_{13}$ equivalent to $\widetilde{\Lambda}_i$. Then in view of the discussion of §5, we may suppose that

$$1 \le R_T \le R_{T-1} \le \ldots \le R_1 \le 5 \quad \text{and} \quad R_1 + \cdots + R_T = 12. \tag{6.4}$$

In addition, since Λ_1 has maximum multiplicity amongst $\Lambda_1, \ldots, \Lambda_{13}$, and multiplicity at most 5, we may suppose that

(a) when none of $\widetilde{\Lambda}_1, \ldots, \widetilde{\Lambda}_T$ are equivalent to Λ_1, then necessarily $T = 12$ and $R_1 = R_2 = \cdots = R_T = 1$, and

(b) when there is a linear form $\widetilde{\Lambda}_j$ equivalent to Λ_1, then necessarily $R_j \le 4$.

Our strategy is to simplify the mean value in question using an argument akin to that employed in the proof of Lemma 6. First, the argument leading to (4.5) above in this instance shows that there is no loss of generality in supposing that

$$\iint_{\mathfrak{E}} |g(\Lambda_1)H_0(\alpha,\beta)|\, d\alpha\, d\beta \ll \iint_{\mathfrak{E}} g_1 \tilde{h}_1^{R_1} \ldots \tilde{h}_T^{R_T}\, d\alpha\, d\beta, \tag{6.5}$$

where here, and in what follows, for each index l we write $\tilde{h}_l$ in place of $|h(\widetilde{\Lambda}_l)|$ and g_l in place of $|g(\Lambda_l)|$. Suppose next that we are in the situation (a) above. We apply Lemma 5 with $N = 4$, with $\tilde{h}_{3l-2}\tilde{h}_{3l-1}\tilde{h}_{3l}$ in place of u_l $(1 \le l \le 4)$, and with u_0 replaced by g_1. By relabelling indices if necessary, we obtain an upper bound of the shape (6.5) in which the exponent sequence $(R_1, \ldots, R_T)$ is equal to $(4, 4, 4)$. Now apply Lemma 5 again with $N = 2$, with $\tilde{h}_l$ in place of u_l $(l = 1, 2)$, and with u_0 replaced by $g_1\tilde{h}_1^3\tilde{h}_2^3\tilde{h}_3^4$. In this way, we conclude that there are indices i, j, k with $1 < i < j < k \le 13$ for which Λ_i, Λ_j and Λ_k are pairwise inequivalent, and

$$\iint_{\mathfrak{E}} |g(\Lambda_1)H_0(\alpha,\beta)|\, d\alpha\, d\beta \ll \iint_{\mathfrak{E}} g_1 h_i^4 h_j^5 h_k^3\, d\alpha\, d\beta. \tag{6.6}$$

We note for future reference the trivial observation that Λ_j is not equivalent to Λ_1.

We analyse the situation (b) by applying an argument paralleling that of the second paragraph of the proof of Lemma 6, in this instance supposing ν to be a non-negative integer for which $R_{T-1} = R_T + \nu < 4$, and now incorporating g_1 into the definition of u_0. Thus, by relabelling indices if necessary, we derive a bound of the shape (6.5), subject to the constraints (6.4) and condition (b) above, wherein $R_{T-1} = 4$. The constraints (6.4) then imply that necessarily $(R_1, R_2, \ldots, R_T) = (5, 4, 3)$ or $(4, 4, 4)$. The latter circumstance may be converted to the former by means of the argument concluding the previous paragraph, and it is apparent that we may ensure in this process that $\widetilde{\Lambda}_2$ remains inequivalent to Λ_1. In this second situation, therefore, we may again conclude that the bound (6.6) holds with $\Lambda_i, \Lambda_j, \Lambda_k$ pairwise inequivalent, and with Λ_j not equivalent to Λ_1.

Define the mean values

$$U = \iint_{\mathfrak{E}} g_1^{21/10} h_i^2 h_j^{77/10}\, d\alpha\, d\beta, \quad V = \iint_{\mathfrak{E}} h_i^8 h_k^8\, d\alpha\, d\beta, \tag{6.7}$$

and, when $(l\,m\,n)$ is a permutation of $(i\,j\,k)$,

$$W_{lmn} = \iint_{\mathfrak{E}} h_l^2 h_m^6 h_n^6\, d\alpha\, d\beta.$$

Then a swift application of Hölder's inequality to (6.6) leads to the bound

$$\iint_{\mathfrak{E}} |g(\Lambda_1) H_0(\alpha,\beta)|\, d\alpha\, d\beta \ll U^{10/21} V^{1/42} W_{ijk}^{3/84} W_{jki}^{35/84} W_{kij}^{1/21}. \tag{6.8}$$

The bound $V = O(P^{10} Q^{-3/100})$ is immediate from Lemma 10(ii), and when Λ_l, Λ_m and Λ_n are pairwise inequivalent, Lemma 7 supplies the estimate $W_{lmn} = O(P^8)$. Thus we conclude from (6.8) that

$$\iint_{\mathfrak{E}} |g(\Lambda_1) H_0(\alpha,\beta)|\, d\alpha\, d\beta \ll P^{89/21} Q^{-1/1400} U^{10/21}. \tag{6.9}$$

It remains to estimate the integral U defined in (6.7). We recall that $\Lambda_1 = a_1\alpha$, and change variables from β to γ via the linear transformation $a_j\alpha + b_j\beta = b_j\gamma$. Note here that since Λ_1 and Λ_j are inequivalent, then necessarily $b_j \neq 0$. Write $A = a_i - b_i a_j / b_j$. Then in view of the definition (6.2) of $\mathfrak{E}$, we may make use of the periodicity of the integrand to deduce that

$$U \le \int_0^1 \int_{\mathfrak{M}} |g(\Lambda_1)|^{21/10} |h(A\alpha + b_i\gamma)|^2 |h(b_j\gamma)|^{77/10}\, d\alpha\, d\gamma \le U_1 U_2, \tag{6.10}$$

where we write

$$U_1 = \int_0^1 |h(b_j\gamma)|^{77/10} d\gamma \quad \text{and} \quad U_2 = \sup_{\lambda\in\mathbb{R}} \int_{\mathfrak{M}} |g(a_1\alpha)|^{21/10} |h(A\alpha+\lambda)|^2\, d\alpha.$$

An application of Lemma 10(i) reveals, via a change of variable, that $U_1 = O(P^{47/10})$, and the bound $U_2 = O(P^{11/10})$ is immediate from Lemma 9. Thus we find from (6.9) and (6.10) that

$$\iint_{\mathfrak{E}} |g(\Lambda_1) H_0(\alpha,\beta)|\, d\alpha\, d\beta \ll P^{89/21} Q^{-1/1400} (P^{29/5})^{10/21} \ll P^7 Q^{-1/1400}.$$

The conclusion of Lemma 8 now follows directly from (6.2), (6.3) and (5.7).

We complete this section with the proofs of Lemmata 9 and 10.

The proof of Lemma 9. Suppose that λ and δ are real numbers with $\delta > 0$. Let A be a fixed non-zero rational number, so that for some $B \in \mathbb{Z} \setminus \{0\}$ and $S \in \mathbb{N}$ with $(B, S) = 1$, one has $A = B/S$. We define the modified set of major arcs $\mathfrak{M}^*$ by putting $\mathfrak{M}^* = \{\beta \in [0, 1) : S\beta \in \mathfrak{M}\}$. Then a change of variable yields the relation

$$\int_{\mathfrak{M}} |g(a_1\theta)|^{2+\delta} |h(A\theta + \lambda)|^2 \, d\theta = S \int_{\mathfrak{M}^*} |g(a_1 S\beta)|^{2+\delta} |h(B\beta + \lambda)|^2 \, d\beta. \qquad (6.11)$$

It follows from the definition of $\mathfrak{M}$ in the preamble to Lemma 7 that for each $\beta \in \mathfrak{M}^*$, there exist $c \in \mathbb{Z}$ and $r \in \mathbb{N}$ with $0 \le c \le r \le P^{3/4}$, $(c, r) = 1$ and $|S\beta r - c| \le P^{-9/4}$. Thus there exist also $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $0 \le a \le q \le SP^{3/4}$, $(a, q) = 1$ and $|q\beta - a| \le P^{-9/4}$. We now take $\kappa(q)$ to be the multiplicative function defined for $q \in \mathbb{N}$ by taking, for primes p and non-negative integers l,

$$\kappa(p^{3l}) = p^{-l}, \quad \kappa(p^{3l+1}) = 2p^{-l-1/2}, \quad \kappa(p^{3l+2}) = p^{-l-1}.$$

Then as a consequence of Theorem 4.1 and Lemmata 4.3 and 4.4 of [19], one has

$$\begin{aligned} g(a_1 S\beta) &\ll \kappa(q) P (1 + P^3 |\beta - a/q|)^{-1} + q^{1/2+\varepsilon} (1 + P^3 |\beta - a/q|)^{1/2} \\ &\ll \kappa(q) P (1 + P^3 |\beta - a/q|)^{-1/2}. \end{aligned}$$

We therefore deduce from (6.11) that

$$\begin{aligned} &\int_{\mathfrak{M}} |g(a_1\theta)|^{2+\delta} |h(A\theta + \lambda)|^2 \, d\theta \\ &\ll \sum_{1 \le q \le SP^{3/4}} (\kappa(q)P)^{2+\delta} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \int_{-\infty}^{\infty} \frac{|h\left(B(a/q + \gamma) + \lambda\right)|^2}{(1 + P^3|\gamma|)^{1+\delta/2}} \, d\gamma. \end{aligned} \qquad (6.12)$$

On making use of the familiar inequality

$$\Big| \sum_{\substack{a=1 \\ (a,q)=1}}^{q} e(al/q) \Big| \le (q, l),$$

we find that

$$\begin{aligned} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |h\left(B\left(a/q + \gamma\right) + \lambda\right)|^2 &= \sum_{x,y \in \mathcal{A}(P,R)} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} e\left((x^3 - y^3)\left(B\left(a/q + \gamma\right) + \lambda\right)\right) \\ &\le |B| \sum_{1 \le x, y \le P} (x^3 - y^3, q). \end{aligned}$$

For each natural number q, write q_0 for the cubefree part of q, and define the integer q_3 via the relation $q = q_0 q_3^3$. Then it follows from the estimate (3.3) of Brüdern and Wooley [7] that whenever $1 \le q \le P$, one has

$$\sum_{1 \le x, y \le P} (x^3 - y^3, q) \ll P^2 q^{\varepsilon} q_3.$$

In this way, we may conclude from (6.12) that

$$\begin{aligned}\int_{\mathfrak{M}} |g(a_1\theta)|^{2+\delta}|h(A\theta+\lambda)|^2\,d\theta \\ &\ll P^{4+\delta}\sum_{1\le q\le SP^{3/4}} q^{\varepsilon}\kappa(q)^{2+\delta}q_3\int_{-\infty}^{\infty}\bigl(1+P^3|\gamma|\bigr)^{-1-\delta/2}\,d\gamma\\ &\ll P^{1+\delta}\sum_{1\le q\le SP^{3/4}} q^{\varepsilon}\kappa(q)^{2+\delta}q_3. \qquad (6.13)\end{aligned}$$

When $\delta>3\varepsilon$, moreover, the sum $\sum_{q=1}^{\infty} q^{\varepsilon}\kappa(q)^{2+\delta}q_3$ converges, as one readily verifies on recalling the definition of $\kappa(q)$. The conclusion of Lemma 9 is now apparent from (6.13).

The proof of Lemma 10. The conclusion of part (i) of Lemma 10 is a special case of Theorem 2 of Brüdern and Wooley [7]. The proof of part (ii) of the lemma requires greater effort. Observe first that from Lemmata 2.2 and 4.4 of [7], it follows easily that when γ is a real number for which $|h(\gamma)|\ge PQ^{-1/10}$, then there exist $a\in\mathbb{Z}$ and $q\in\mathbb{N}$ with $1\le q\le Q^{1/3}$, $(a,q)=1$ and $|q\alpha-a|\le Q^{1/3}P^{-3}$. Consequently, if Λ_k and Λ_l are inequivalent linear forms and $h_kh_l\ge P^2Q^{-1/10}$, then for $\sigma=k,l$ there exist integers d_σ and q_σ with

$$1\le q_\sigma\le Q^{1/3},\quad (d_\sigma,q_\sigma)=1\quad\text{and}\quad |\Lambda_\sigma-d_\sigma/q_\sigma|\le q_\sigma^{-1}Q^{1/3}P^{-3}.$$

Write $c=a_lb_k-a_kb_l$ and $\tau_c=c/|c|$, and consider the linear expressions $a_l\Lambda_k-a_k\Lambda_l$ and $b_k\Lambda_l-b_l\Lambda_k$. Then we see that in the circumstances at hand, one has $(\alpha,\beta)\in\mathfrak{N}(q,a,b)$, where

$$q=|c|q_lq_k,\quad a=\tau_c(b_kd_lq_k-b_ld_kq_l)\quad\text{and}\quad b=\tau_c(a_ld_kq_l-a_kd_lq_k).$$

It follows inter alia that when $h_kh_l\ge P^2Q^{-1/10}$, one necessarily has $(\alpha,\beta)\in\mathfrak{N}$. We therefore deduce that

$$\sup_{(\alpha,\beta)\in\mathfrak{n}}(h_kh_l)\ll P^2Q^{-1/10},$$

whence

$$\iint_{\mathfrak{n}} h_k^8h_l^8\,d\alpha\,d\beta\ll (P^2Q^{-1/10})^{3/10}\int_0^1\int_0^1 (h_kh_l)^{77/10}\,d\alpha\,d\beta. \qquad (6.14)$$

On making use of the first conclusion of the lemma in combination with a change of variables, one finds that

$$\int_0^1\int_0^1(h_kh_l)^{77/10}\,d\alpha\,d\beta\ll\biggl(\int_0^1|h(\xi)|^{77/10}\,d\xi\biggr)^2\ll P^{47/5},$$

and so the conclusion of the second part of the lemma follows from (6.14).

7. The major arc analysis. We now turn our attention to the problem of estimating the contribution to the integral in (5.6) that arises from the major arcs $\mathfrak{N}$. There are relatively few variables involved in this integral, and our current set-up avoids various artifices that earlier writers have employed. For these reasons, there is no suitable reference available in the literature. However, the argument that we apply is nonetheless largely standard, and so we shall be brief.

First we introduce the approximants to the generating functions g and h on the major arcs $\mathfrak{N}$. Let

$$S(q,r)=\sum_{l=1}^{q} e(rl^3/q) \quad \text{and} \quad S_i(q,c,d)=S(q,a_ic+b_id) \quad (1\le i\le s).$$

Also, write

$$v(\theta)=\int_0^P e(\theta\gamma^3)\,d\gamma \quad \text{and} \quad w(\theta)=\int_{\eta P}^P e(\theta\gamma^3)\,d\gamma. \tag{7.1}$$

Finally, we mimic the convention (5.1) by associating with the pair (a_j,b_j) the linear form $\lambda_j=a_j\xi+b_j\zeta$ for $1\le j\le s$, and when it is convenient for the task at hand, we write also $v_j(\xi,\zeta)=v(\lambda_j)$. From Lemma 8.5 of [20] (see also Lemma 5.4 of [18]), it follows that there exists a positive number ρ, depending at most on η, such that whenever $(\alpha,\beta)\in\mathfrak{N}(q,a,b)\subseteq\mathfrak{N}$, then

$$h(\Lambda_j)-\rho q^{-1}S_i(q,a,b)v_i(\alpha-a/q,\beta-b/q)\ll P(\log P)^{-1/2}. \tag{7.2}$$

Similarly, as a consequence of Theorem 4.1 of [19], one finds that under the same constraints on (α,β), one has

$$g(\Lambda_1)-q^{-1}S_1(q,a,b)w(a_1(\alpha-a/q))\ll \log P. \tag{7.3}$$

Here we have made use of the hypothesis, justified by the discussion of section 5 and recorded in (5.3), that $b_1=0$, whence in particular $\Lambda_1=a_1\alpha$. On writing

$$V(\xi,\zeta)=w(a_1\xi)\prod_{j=2}^{s} v(\lambda_j) \quad \text{and} \quad U(q,a,b)=q^{-s}\prod_{j=1}^{s} S_i(q,a,b), \tag{7.4}$$

and recalling the definition (5.5), we deduce from (7.2) and (7.3) that the estimate

$$g(\Lambda_1)H(\alpha,\beta)-\rho^{s-1}U(q,a,b)V(\alpha-a/q,\beta-b/q)\ll P^s(\log P)^{-1/2} \tag{7.5}$$

holds whenever $(\alpha,\beta)\in\mathfrak{N}(q,a,b)\subseteq\mathfrak{N}$.

Next we introduce truncated versions of the singular integral and singular series, which we define respectively by

$$\mathfrak{J}(X)=\iint_{\mathfrak{B}(X)} V(\xi,\zeta)\,d\xi\,d\zeta \quad \text{and} \quad \mathfrak{S}(X)=\sum_{1\le q\le X} A(q), \tag{7.6}$$

in which we have written $\mathfrak{B}(X)$ for the box $[-XP^{-3}, XP^{-3}]^2$, and where

$$A(q) = \sum_{\substack{c=1 \\ (c,d,q)=1}}^{q} \sum_{d=1}^{q} U(q,c,d). \tag{7.7}$$

The measure of the major arcs $\mathfrak{N}$ is $O(Q^5P^{-6})$, so that on recalling (5.7) and integrating over $\mathfrak{N}$, we infer from (7.5) that

$$\iint_{\mathfrak{N}} g(\Lambda_1)H(\alpha,\beta)\,d\alpha\,d\beta - \rho^{s-1}\mathfrak{S}(Q)\mathfrak{J}(Q) \ll P^{s-6}(\log P)^{-1/4}. \tag{7.8}$$

It now remains only to analyse the singular series and the singular integral defined, in truncated form, in (7.6). With an application in our forthcoming article [8] in mind, we study $\mathfrak{S}(X)$ and $\mathfrak{J}(X)$ in a slightly more general situation than is warranted for the application at hand, and suppose only that for any $(c,d) \in \mathbb{Z}^2 \setminus \{(0,0)\}$, at least seven of the numbers $ca_j + db_j$ $(1 \le j \le s)$ are non-zero. In this new more general context, it is possible that a given linear form Λ_i may have multiplicity as high as six from amongst $\Lambda_1, \dots, \Lambda_{13}$. Fortunately, the proofs of Lemmata 12 and 13 below would be no simpler if this additional case were to be excluded.

In preparation for our discussion of the singular series, we introduce some additional notation and provide a simple auxiliary estimate. When $1 \le j \le 13$ and $(c,d) \in \mathbb{Z}^2 \setminus \{0,0\}$, we define the integer $u_j = u_j(c,d)$ by

$$u_j = (q, ca_j + db_j). \tag{7.9}$$

We suppose that $\{j_1, \dots, j_t\} \subseteq \{1, \dots, 13\}$ is a maximal set of distinct subscripts with the property that the linear forms Λ_{j_k} are pairwise inequivalent for $1 \le k \le t$. It is convenient then to define the integers

$$\Delta = \prod_{1 \le k < l \le t} |a_{j_k}b_{j_l} - a_{j_l}b_{j_k}| \quad \text{and} \quad \Xi = \Delta^2 \prod_{1 \le k \le t} (a_{j_k}, b_{j_k}). \tag{7.10}$$

Lemma 11. *When $q \in \mathbb{N}$ and $(c,d) \in \mathbb{Z}^2$ satisfy the condition $(q,c,d) = 1$, one has $u_{j_1}u_{j_2}\dots u_{j_t} | \Delta q$. Moreover, when $v_1, \dots, v_t$ are integers with $v_1v_2\dots v_t | \Delta q$, there are at most $\Xi q^2(v_1 \dots v_t)^{-1}$ integral pairs (c,d) with $1 \le c,d \le q$ satisfying $(c,d,q) = 1$ and $u_{j_k} = v_k$ $(1 \le k \le t)$.*

Proof. Although the desired conclusions may be extracted from the argument of the proof of Lemma 35 of Davenport and Lewis [12], we provide a brief proof here for the sake of transparency of exposition. Suppose first that $q \in \mathbb{N}$ and $(c,d) \in \mathbb{Z}^2$ satisfy $(q,c,d) = 1$. By manipulating appropriate linear combinations of arguments, one sees that for $1 \le k < l \le t$ one has

$$(q, ca_{j_l} + db_{j_l}, ca_{j_k} + db_{j_k}) | (q, a_{j_l}b_{j_k} - a_{j_k}b_{j_l})(q,c,d).$$

By hypothesis, we may suppose that $a_{j_l}b_{j_k} \neq a_{j_k}b_{j_l}$, and thus we deduce from (7.9) and (7.10) that

$$\prod_{1\le k<l\le t} (u_{j_k}, u_{j_l})|\Delta. \tag{7.11}$$

The desired conclusion $u_{j_1}u_{j_2}\dots u_{j_t}|\Delta q$ then follows from the observation that $u_{j_l}|q$ for $1\le l\le t$. Next we note that for $1\le l\le t$, the number of solutions (c,d), distinct modulo v_l, of the congruence $a_{j_l}c + b_{j_l}d \equiv 0 \pmod{v_l}$, is precisely $(a_{j_l}, b_{j_l}, v_l)v_l$. On recalling (7.9) and applying the Chinese Remainder Theorem, therefore, the number of integral pairs (c,d) satisfying $1\le c,d\le q$, $(q,c,d)=1$ and $u_{j_l}=v_l$ $(1\le l\le t)$ is at most

$$\left(\prod_{1\le l\le t} (a_{j_l}, b_{j_l})v_l\right)\left(q(v_1\dots v_t)^{-1}\prod_{1\le k<l\le t}(v_k,v_l)\right)^2.$$

Now $(v_k,v_l)=(u_{j_k},u_{j_l})$, so on making use of (7.11) in order to bound the last product in this expression, we conclude from (7.10) that an upper bound for the number of integral pairs (c,d) in question is $\Xi q^2(v_1\dots v_t)^{-1}$. This confirms the final conclusion of the lemma.

As will shortly be confirmed, the singular series $\mathfrak{S}$ is equal to the product of the p-adic densities of solutions. In this context we define the p-adic density χ_p by

$$\chi_p = \lim_{h\to\infty} p^{h(2-s)}M_s(p^h), \tag{7.12}$$

where we write $M_s(p^h)$ for the number of solutions of the system (1.1) with $\mathbf{x}\in(\mathbb{Z}/p^h\mathbb{Z})^s$.

Lemma 12. *Suppose that the linear forms $L_1(\boldsymbol{\theta})$ and $L_2(\boldsymbol{\theta})$ associated with the system (1.1) satisfy the condition that for any pair $(c,d)\in\mathbb{Z}^2\setminus\{(0,0)\}$, the linear form $cL_1(\boldsymbol{\theta})+dL_2(\boldsymbol{\theta})$ contains at least seven non-zero coefficients. Then the limit $\mathfrak{S}=\lim_{X\to\infty}\mathfrak{S}(X)$ exists, and one has*

$$\mathfrak{S}-\mathfrak{S}(X)\ll X^{-1/4}. \tag{7.13}$$

Moreover, the Euler product $\prod_p \chi_p$ converges absolutely to $\mathfrak{S}$, and one has $\chi_p=0$ if and only if the system (1.1) has no non-trivial solution in $\mathbb{Q}_p$. Finally, when the system (1.1) possesses a non-trivial solution in $\mathbb{Q}_p$ for every prime number p, one has $\mathfrak{S}\gg 1$.

Proof. We establish the upper bound

$$A(q)\ll q^{\varepsilon-4/3}. \tag{7.14}$$

In view of (7.6), the estimate (7.14) not only confirms (7.13) but also shows that the limit $\mathfrak{S}=\lim_{X\to\infty}\mathfrak{S}(X)$ exists. The proof of the remaining conclusions of the lemma follow by the theory familiar to practitioners of the circle method (see, for example, Section 2.6 of [19], or Section 10 of Davenport and Lewis [12]). From (7.7) and (7.14) one finds that

$$p^{h(2-s)}M_s(p^h)=\sum_{l=1}^{p^h}A(p^l)=1+O(p^{\varepsilon-4/3}),$$

so that the definition (7.12) yields the estimate $\chi_p = 1+O(p^{-5/4})$. It follows that the Euler product $\prod_p \chi_p$ converges absolutely. But from (7.7) we see that $A(q)$ is a multiplicative function of q, and so we see from (7.6) that indeed $\mathfrak{S}$ is equal to the aforementioned Euler product. We may choose p_0 large enough so that $\chi_p \ge 1-p^{-6/5}$ for $p>p_0$, and then it follows that

$$\mathfrak{S} \gg \prod_{p\le p_0}\chi_p. \tag{7.15}$$

In circumstances wherein the system (1.1) has no non-trivial solution in $\mathbb{Q}_p$, it fails to possess a non-trivial solution in $\mathbb{Z}/p^h\mathbb{Z}$ for any h, and thus it follows from (7.12) that $\chi_p = 0$. On the other hand, when the system (1.1) possesses a non-trivial solution in $\mathbb{Q}_p$, the argument of the proof of the Corollary to Theorem 1 of Davenport and Lewis [12] (see the end of section 5 of the latter paper) shows that the system (1.1) has a non-singular solution in $\mathbb{Q}_p$. An argument employing Hensel's Lemma (as in Lemma 6.7 of [20], for example) then shows that $M_s(p^h)\gg p^{h(2-s)}$ for large enough values of h, whence (7.12) shows that $\chi_p>0$. Thus $\chi_p=0$ if and only if the system (1.1) has no non-trivial solution in $\mathbb{Q}_p$, and by (7.15) one has $\mathfrak{S}\gg 1$ unless the system (1.1) fails to possess a non-trivial solution in $\mathbb{Q}_p$ for some prime p.

It remains to establish (7.14). First, by relabelling indices if necessary, the hypotheses of the lemma permit the assumption that the maximum multiplicity of any of the forms $\Lambda_1,\dots,\Lambda_{13}$ is six. By Theorem 4.2 of [19], whenever $(q,r)=1$ one has $S(q,r)\ll q^{2/3}$. Thus, on recalling the definition (7.9), one finds that

$$q^{-1}S(q,ca_j+db_j)\ll u_j^{1/3}q^{-1/3}.$$

Consequently, using trivial estimates for factors in the definition (7.4) of $U(q,c,d)$ with $j>13$, we deduce that

$$U(q,c,d)\ll q^{-13/3}(u_1u_2\dots u_{13})^{1/3}.$$

We note now that when Λ_l and Λ_k are equivalent linear forms, it is a consequence of (7.9) that $u_l\ll u_k\ll u_l$. Recall the notation defined in the preamble to Lemma 11, and suppose that for $1\le k\le t$, the linear form Λ_{j_k} has multiplicity r_k amongst

$\Lambda_1, \dots, \Lambda_{13}$. Then in view of (7.7), it follows from Lemma 11 that

$$\begin{aligned} A(q) &\ll q^{-13/3} \sum_{\substack{c=1 \\ (c,d,q)=1}}^{q} \sum_{d=1}^{q} (u_1 u_2 \dots u_{13})^{1/3} \\ &\ll q^{-13/3} \sum_{\substack{v_1,\dots,v_t \\ v_1\dots v_t | \Delta q}} \frac{\Xi q^2}{v_1 \dots v_t} (v_1^{r_1} \dots v_t^{r_t})^{1/3}. \end{aligned}$$

Observe that since the linear forms $\Lambda_{j_1}, \dots, \Lambda_{j_t}$ are pairwise inequivalent, the integer Δ is non-zero, and further, the integers Δ and Ξ are bounded purely in terms of the coefficients $\mathbf{a}$ and $\mathbf{b}$. We are permitted to assume that $r_j \le 6$ for $1 \le j \le t$, so on using an elementary bound for the divisor function, we conclude that

$$A(q) \ll q^{-7/3} \sum_{\substack{v_1,\dots,v_t \\ v_1\dots v_t | \Delta q}} v_1 \dots v_t \ll q^{\varepsilon - 4/3},$$

as claimed in (7.14). This completes the proof of the lemma.

We now turn our attention to the truncated singular integral $\mathfrak{J}(X)$. The analysis here is very straightforward, but ironically, the simplicity of our approach prevents any convenient reference to the literature.

Lemma 13. *Under the same hypotheses as in the statement of Lemma* 12*, the limit* $\mathfrak{J} = \lim_{X \to \infty} \mathfrak{J}(X)$ *exists, and one has*

$$\mathfrak{J} - \mathfrak{J}(X) \ll P^{s-6} X^{-1}. \tag{7.16}$$

Moreover, one has $\mathfrak{J} \gg P^{s-6}$.

Proof. We begin by considering two inequivalent forms Λ_i and Λ_j. When T is a positive number, write $\mathfrak{B}(T)$ for the box $[-T, T]^2$, and $\widehat{\mathfrak{B}}(T)$ for the complementary set $\mathbb{R}^2 \backslash \mathfrak{B}(T)$. Consider now a positive number Y and suppose that $(\Lambda_i, \Lambda_j) \in \mathfrak{B}(Y)$. The latter assertion is equivalent to the statement that

$$a_i \alpha + b_i \beta \in [-Y, Y] \quad \text{and} \quad a_j \alpha + b_j \beta \in [-Y, Y],$$

so that on taking suitable linear combinations of these forms, one obtains

$$(a_i b_j - a_j b_i)\alpha \in |b_j|\,[-Y, Y] + |b_i|\,[-Y, Y]$$

and

$$(a_j b_i - a_i b_j)\beta \in |a_j|\,[-Y, Y] + |a_i|\,[-Y, Y],$$

the sums of intervals being interpreted set theoretically. The integer $a_i b_j - a_j b_i$ is non-zero, so that if we write

$$\Theta^{-1} = \max_{1 \le i < j \le s} \{|a_i| + |a_j|\,,\, |b_i| + |b_j|\},$$

then we may conclude that $(\alpha,\beta)\in\mathfrak{B}(\Theta^{-1}Y)$. On making a transparent change of variables, it follows from this discussion that when Λ_i and Λ_j are inequivalent, one has

$$\iint_{\widehat{\mathfrak{B}}(XP^{-3})}|v(\lambda_i)v(\lambda_j)|^6\,d\xi\,d\zeta\ll\iint_{\widehat{\mathfrak{B}}(\Theta XP^{-3})}|v(\lambda_i)v(\lambda_j)|^6\,d\lambda_i\,d\lambda_j. \tag{7.17}$$

We now recall the estimate $v(\gamma)\ll P(1+P^3|\gamma|)^{-1/3}$ that follows, for example, from Theorem 7.3 of [19]. In the situation at hand, we may suppose that none of the forms $\Lambda_1,\dots,\Lambda_{13}$ has multiplicity exceeding six. Hence, following a suitable relabelling of indices, we may temporarily suppose that for $1\le k\le 6$, the forms Λ_{2k}, Λ_{2k+1} are inequivalent. Then on integrating the elementary inequality

$$|v(\lambda_2)v(\lambda_3)\dots v(\lambda_{13})|\le\sum_{k=1}^{6}|v(\lambda_{2k})v(\lambda_{2k+1})|^6$$

and applying the observation (7.17), we deduce that

$$\begin{aligned}\iint_{\widehat{\mathfrak{B}}(XP^{-3})}&|v(\lambda_2)v(\lambda_3)\dots v(\lambda_{13})|\,d\xi\,d\zeta\\ &\ll P^{12}\iint_{\widehat{\mathfrak{B}}(\Theta XP^{-3})}(1+P^3|\lambda|)^{-2}(1+P^3|\mu|)^{-2}\,d\lambda\,d\mu\\ &\ll P^6X^{-1}.\end{aligned}$$

Finally, on using trivial bounds for $w(a_1\xi)$ and $v(\lambda_j)$ when $j>13$, we conclude that

$$\iint_{\widehat{\mathfrak{B}}(XP^{-3})}|V(\xi,\zeta)|\,d\xi\,d\zeta\ll P^{s-6}X^{-1}. \tag{7.18}$$

In particular, it follows from (7.6) that the singular integral $\mathfrak{J}=\lim\limits_{X\to\infty}\mathfrak{J}(X)$ exists, and (7.18) provides the desired estimate (7.16).

In order to evaluate $\mathfrak{J}$ we follow the familiar routine based on the use of Fourier's integral theorem. From (7.1) and (7.6), we see that

$$\mathfrak{J}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{\mathfrak{D}}e(\xi\mathcal{L}_1(\boldsymbol{\gamma})+\zeta\mathcal{L}_2(\boldsymbol{\gamma}))\,d\boldsymbol{\gamma}\,d\xi\,d\zeta,$$

where we write

$$\mathcal{L}_1(\boldsymbol{\gamma})=\sum_{i=1}^{s}a_i\gamma_i^3\quad\text{and}\quad\mathcal{L}_2(\boldsymbol{\gamma})=\sum_{i=1}^{s}b_i\gamma_i^3,$$

and where $\mathcal{D}$ denotes the box $[\eta P,P]\times[0,P]^{s-1}$. Put $\mu=P^3\xi$ and $\nu=P^3\zeta$, and substitute $\omega_i=(P^{-1}\gamma_i)^3$ for $1\le i\le s$. Then with these changes of variables we discover that

$$\mathfrak{J}=3^{-s}P^{s-6}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{\mathfrak{D}'}\frac{e(\mu L_1(\boldsymbol{\omega})+\nu L_2(\boldsymbol{\omega}))}{(\omega_1\dots\omega_s)^{2/3}}\,d\boldsymbol{\omega}\,d\mu\,d\nu,$$

where $L_1(\boldsymbol{\omega})$ and $L_2(\boldsymbol{\omega})$ are defined in (5.2), and where $\mathfrak{D}' = [\eta^3, 1] \times [0,1]^{s-1}$. The discussion of section 5 ensures that the equations $L_1(\boldsymbol{\omega}) = L_2(\boldsymbol{\omega}) = 0$ define an $(s-2)$-dimensional linear space that passes through the point $\boldsymbol{\theta}$ lying in the interior of $\mathfrak{D}'$. Recall from (5.3) that $b_1 = a_2 = 0$, whence from (5.2),

$$\omega_1 = a_1^{-1}\Big(L_1(\boldsymbol{\omega}) - \sum_{j=3}^{s} a_j\omega_j\Big) \quad \text{and} \quad \omega_2 = b_2^{-1}\Big(L_2(\boldsymbol{\omega}) - \sum_{j=3}^{s} b_j\omega_j\Big).$$

Then on making a change of variables and applying Fourier's integral formula twice, in the shape

$$\lim_{\lambda\to\infty}\int_{-T}^{T}\int_{-\lambda}^{\lambda} V(t)e(t\omega)\,d\omega\,dt = V(0),$$

we obtain the relation

$$\mathfrak{J} \gg P^{s-6}\int_{\mathfrak{D}''}(\omega_1\dots\omega_s)^{-2/3}\,d\omega_3\,d\omega_4\dots d\omega_s. \tag{7.19}$$

Here, we define the coordinates ω_1 and ω_2 by

$$\omega_1 = -a_1^{-1}\sum_{j=3}^{s} a_j\omega_j \quad \text{and} \quad \omega_2 = -b_2^{-1}\sum_{j=3}^{s} b_j\omega_j,$$

and we write $\mathfrak{D}''$ for the set of $(s-2)$-tuples $(\omega_3,\omega_4,\dots,\omega_s) \in [0,1]^{s-2}$ for which the s-tuple $(\omega_1,\dots,\omega_s)$ lies in $\mathfrak{D}'$. Notice that the point $(\theta_3,\dots,\theta_s)$ necessarily lies in the interior of the polytope $\mathfrak{D}''$, whence $\mathfrak{D}''$ has positive volume. The latter observation ensures that the integral on the right hand side of (7.19) is positive. Since, plainly, the latter integral is independent of P, we may conclude that $\mathfrak{J} \gg P^{s-6}$, and this completes the proof of the lemma.

The proof of Theorem 2 is now swiftly completed. By (7.8) and Lemmata 12 and 13, we find that

$$\iint_{\mathfrak{N}} g(\Lambda_1)H(\alpha,\beta)\,d\alpha\,d\beta - \rho^{s-1}\mathfrak{S}\mathfrak{J} \ll P^{s-6}Q^{-1/4},$$

so that in view of the estimate (5.9) we may conclude that

$$\int_0^1\int_0^1 g(\Lambda_1)H(\alpha,\beta)\,d\alpha\,d\beta = \rho^{s-1}\mathfrak{S}\mathfrak{J} + O(P^{s-6}(\log P)^{-1/140000}).$$

From Lemmata 12 and 13, moreover, it is apparent that $\rho^{s-1}\mathfrak{S}\mathfrak{J} \gg P^{s-6}$ provided only that the system (1.1) has non-trivial solutions in $\mathbb{Q}_p$ for every prime p. But in such circumstances, the lower bound (5.6) ensures that $\mathcal{N}(P) \gg P^{s-6}$. In view of the discussion on p-adic solubility prior to the statement of Theorem 2, solubility over $\mathbb{Q}_p$ is already assured when $p \neq 7$, and the conclusion of Theorem 2 follows immediately.

8. Le coup de grâce. The theme of this concluding section is the proof of Theorem 1. Needless to say, if Theorem 2 is applicable to the system (1.1), then there is nothing further to discuss. Thus we may suppose that there exists a pair $(c,d) \in \mathbb{Z}^2 \setminus \{(0,0)\}$ with the property that at most seven of the numbers $ca_j + db_j$ $(1 \le j \le s)$ are non-zero. By taking suitable rational linear combinations of the two equations defining (1.1), it is apparent that there is no loss of generality in supposing that the system (1.1) takes the shape

$$a_1x_1^3 + \ldots + a_sx_s^3 = b_{t+1}x_{t+1}^3 + \ldots + b_sx_s^3 = 0, \tag{8.1}$$

where $s \ge 13$ and $s - t \le 7$. We now recall a conclusion of R. Baker concerning the solubility of diagonal cubic equations.

Lemma 14. *Whenever $r \ge 7$ and $c_1, \ldots, c_r$ are rational integers, the equation $c_1x_1^3 + \ldots + c_rx_r^3 = 0$ possesses a non-trivial integral solution.*

Proof. On setting $x_i = 0$ for $i > 7$, the desired conclusion follows from [1].

Let us return to the system (8.1). If one has $s - t \le 6$, then the condition $s \ge 13$ implies that $t \ge 7$, and indeed the same conclusion holds when $s \ge 14$ and $s - t = 7$. But when $t \ge 7$, it follows from Lemma 14 that the equation $a_1y_1^3 + \ldots + a_ty_t^3 = 0$ possesses a non-trivial integral solution $\mathbf{y} = \mathbf{z}$, and thus the system (8.1) has the non-trivial integral solution $\mathbf{x} = (z_1, \ldots, z_t, 0, \ldots, 0)$. We are therefore left to ponder the situation in which $s - t = 7$ and $s = 13$. In view of Lemma 14, the equation $b_7y_7^3 + \ldots + b_{13}y_{13}^3 = 0$ possesses a non-trivial integral solution $(y_7, \ldots, y_{13}) = (z_7, \ldots, z_{13})$. We put $A = a_7z_7^3 + \cdots + a_{13}z_{13}^3$, and consider the equation $Ay_0^3 + a_1y_1^3 + \ldots + a_6y_6^3 = 0$. This equation possesses a non-trivial solution $\mathbf{y} = \mathbf{w}$, again by Lemma 14, and so the system (8.1) in this instance has the non-trivial integral solution

$$(x_1, \ldots, x_{13}) = (w_1, \ldots, w_6, w_0z_7, \ldots, w_0z_{13}).$$

We therefore conclude that when $s \ge 13$ and Theorem 2 fails to deliver the Hasse principle for the system (1.1), this system nonetheless possesses non-trivial integral solutions. This completes the proof of Theorem 1.

References

1. R. C. Baker, *Diagonal cubic equations II*, Acta Arith. **53** (1989), 217–250.
2. R. C. Baker and J. Brüdern, *On pairs of additive cubic equations*, J. Reine Angew. Math. **391** (1988), 157–180.
3. B. J. Birch, *Homogeneous forms of odd degree in a large number of variables*, Mathematika **4** (1957), 102–105.
4. B. J. Birch, *Forms in many variables*, Proc. Roy. Soc. Ser. A **265** (1961), 245–263.
5. J. Brüdern, *On pairs of diagonal cubic forms*, Proc. London Math. Soc. (3) **61** (1990), 273–343.
6. J. Brüdern, K. Kawada and T. D. Wooley, *Additive representation in thin sequences, I: Waring's problem for cubes*, Ann. Sci. École Norm. Sup. (4) **34** (2001), 471–501.

7. J. Brüdern and T. D. Wooley, *On Waring's problem for cubes and smooth Weyl sums*, Proc. London Math. Soc. (3) **82** (2001), 89–109.
8. J. Brüdern and T. D. Wooley, *The density of integral solutions for pairs of diagonal cubic equations* (preprint).
9. J. W. S. Cassels and M. J. T. Guy, *On the Hasse principle for cubic surfaces*, Mathematika **13** (1966), 111–120.
10. R. J. Cook, *Pairs of additive equations*, Michigan Math. J. **19** (1972), 325–331.
11. R. J. Cook, *Pairs of additive congruences: cubic congruences*, Mathematika **32** (1985), 286–300.
12. H. Davenport and D. J. Lewis, *Cubic equations of additive type*, Philos. Trans. Roy. Soc. London Ser. A **261** (1966), 97–136.
13. R. Dietmann and T. D. Wooley, *Pairs of cubic forms in many variables*, Acta Arith. **110** (2003), 125–140.
14. D. J. Lewis, *Cubic forms over algebraic number fields*, Mathematika **4** (1957), 97–101.
15. H. P. F. Swinnerton-Dyer, *The solubility of diagonal cubic surfaces*, Ann. Sci. École Norm. Sup. (4) **34** (2001), 891–912.
16. R. C. Vaughan, *On pairs of additive cubic equations*, Proc. London Math. Soc. (3) **34** (1977), 354–364.
17. R. C. Vaughan, *On Waring's problem for cubes*, J. Reine Angew. Math. **365** (1986), 122–170.
18. R. C. Vaughan, *A new iterative method in Waring's problem*, Acta Math. **162** (1989), 1–71.
19. R. C. Vaughan, *The Hardy-Littlewood Method, second edition*, Cambridge University Press, 1997.
20. T. D. Wooley, *On simultaneous additive equations, II*, J. Reine Angew. Math. **419** (1991), 141–198.
21. T. D. Wooley, *Breaking classical convexity in Waring's problem: sums of cubes and quasi-diagonal behaviour*, Invent. Math. **122** (1995), 421–451.
22. T. D. Wooley, *Sums of three cubes*, Mathematika **47** (2000), 53–61.
23. T. D. Wooley, *Slim exceptional sets for sums of cubes*, Canad. J. Math. **54** (2002), 417–448.
24. T. D. Wooley, *Slim exceptional sets and the asymptotic formula in Waring's problem*, Math. Proc. Cambridge Philos. Soc. **134** (2003), 193–206.
25. T. D. Wooley, *Slim exceptional sets for stout representation problems* (preprint).

Jörg Brüdern: Institut für Algebra und Zahlentheorie, Pfaffenwaldring 57, Universität Stuttgart, D-70511 Stuttgart, Germany
E-mail address: `bruedern@mathematik.uni-stuttgart.de`

Trevor D. Wooley: Department of Mathematics, University of Michigan, 2074 East Hall, 530 Church Street, Ann Arbor, MI 48109-1043, U.S.A.
E-mail address: `wooley@umich.edu`