PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The addition of binary cubic forms

Jörg Brüdern and Trevor D. Wooley

Phil. Trans. R. Soc. Lond. A 1998 **356**, 701-737
doi: 10.1098/rsta.1998.0182

The Royal Society

The addition of binary cubic forms

By Jörg Brüdern[1] and Trevor D. Wooley[2]

[1]*Mathematisches Institut A, Universität Stuttgart, Postfach 80 11 40, D-70511 Stuttgart, Germany*
[2]*Department of Mathematics, University of Michigan, East Hall, 525 East University Avenue, Ann Arbor, MI 48109-1109, USA*

We show that a sum of four non-degenerate binary cubic forms with integral coefficients necessarily possesses a non-trivial rational zero. When each of these binary cubic forms has non-zero discriminant, we are able to obtain bounds on the number, $\mathcal{N}(P)$, of integral zeros of the sum inside a box of size P of the shape

$$P^{5-\varepsilon} \ll_\varepsilon \mathcal{N}(P) \ll_\varepsilon P^{5+\varepsilon}.$$

Finally, given two binary cubic forms with non-zero discriminant, we show that almost all integers, lying in those congruence classes permitted by local solubility conditions, are represented as the sum of the aforementioned forms.

Keywords: Hardy–Littlewood method; exponential sums; diophantine equations; cubic forms; representation problems; efficient differencing

1. Introduction

The circle method, originally designed by Hardy and Littlewood for application to Waring's problem, has also been wrought profitably in investigations of the solubility of non-diagonal diophantine equations, with admirable success in the particular case of cubic equations. In recent years Hooley (1988, 1991, 1994) has considered non-singular cubic forms Φ, with integer coefficients, in $s \geqslant 9$ variables, and has shewn that the equation $\Phi = 0$ has a non-trivial integer solution (that is, a solution $\boldsymbol{x} \neq \mathbf{0}$) if and only if Φ has non-singular zeros in every p-adic field. When $s \geqslant 10$ one knows that non-singular p-adic solutions always exist, and in such circumstances we may therefore conclude that the variety defined by the equation $\Phi = 0$ must have rational points; this much was established earlier by Heath-Brown (1983). One can avoid non-singularity conditions if one is prepared to accept more variables. In a pioneering series of papers, Davenport (1959, 1962, 1963) unconditionally established the existence of non-trivial rational zeros for all cubic forms in 16 or more variables (see also Hooley (1991) when the singularities are mild). If the form is diagonal, the equation takes the shape

$$a_1x_1^3 + a_2x_2^3 + \cdots + a_sx_s^3 = 0,$$

with integer coefficients a_i $(1 \leqslant i \leqslant s)$. As a consequence of the work of Baker (1989), it is known that the latter equation possesses non-trivial solutions whenever $s \geqslant 7$.

The question now arises as to whether one can break away from the diagonal situation when the number of variables does not exceed 8. A first attempt was made by Chowla & Davenport (1961) over three decades ago. They considered binary cubic

forms $\Phi_j \in \mathbb{Z}[x,y]$ $(j = 1,2,3)$ with non-zero discriminant, and showed the existence of a non-trivial solution of the diophantine equation

$$\Phi_1(x_1,y_1) + \Phi_2(x_2,y_2) + \Phi_3(x_3,y_3) + ax_4^3 + by_4^3 = 0,$$

where $a, b \in \mathbb{Z}$. Their method, however, makes crucial use of the presence of the two isolated variables, x_4 and y_4 (see lemma 6 of Chowla & Davenport (1961)), and does not extend to handle the situation in which the diagonal form $ax_4^3 + by_4^3$ is replaced by a general non-singular binary cubic form. Our primary objective in this paper is to present a method which on the one hand deals with the addition of four binary cubic forms in full generality, and on the other hand permits more control to be exercised on the number of solutions.

Theorem 1.1. *Let $\Phi_j \in \mathbb{Z}[x,y]$ $(1 \leqslant j \leqslant 4)$ be binary cubic forms with integer coefficients and non-zero discriminants. Let $\mathcal{N}(P) = \mathcal{N}(P;\boldsymbol{\Phi})$ denote the number of solutions of the diophantine equation*

$$\Phi_1(x_1,y_1) + \Phi_2(x_2,y_2) + \Phi_3(x_3,y_3) + \Phi_4(x_4,y_4) = 0, \tag{1.1}$$

subject to $|x_j| \leqslant P$ and $|y_j| \leqslant P$ $(1 \leqslant j \leqslant 4)$. Then for each $\varepsilon > 0$ one has

$$P^{5-\varepsilon} \ll_{\varepsilon,\boldsymbol{\Phi}} \mathcal{N}(P;\boldsymbol{\Phi}) \ll_{\varepsilon,\boldsymbol{\Phi}} P^{5+\varepsilon}.$$

If Φ is a binary cubic form with integer coefficients and zero discriminant, then Φ may be transformed by a rational change of variables to one or other of u^3 and uv^2. An equation of the type (1.1), in which one or more of the Φ_i has zero discriminant, is therefore immediately soluble non-trivially.

As experts in the field will immediately recognize, one should expect a cognate result on the addition of two binary cubic forms.

Theorem 1.2. *Let $\Phi_j \in \mathbb{Z}[x,y]$ $(j = 1,2)$ be binary cubic forms with non-zero discriminant. Let $\mathcal{W}$ denote the set of all positive integers, n, for which the congruence*

$$\Phi_1(x_1,y_1) + \Phi_2(x_2,y_2) \equiv n \pmod{q} \tag{1.2}$$

has a solution for all $q \in \mathbb{N}$. Then the set $\mathcal{W}$ has positive density, and the number $E(N)$ of all natural numbers $n \in \mathcal{W}$ not exceeding N, for which the diophantine equation

$$\Phi_1(x_1,y_1) + \Phi_2(x_2,y_2) = n \tag{1.3}$$

has no solution, satisfies $E(N) \ll N^{209/210+\varepsilon}$.

The conclusion of theorem 1.2 should be compared with a similar result in Chowla & Davenport (1961) (see the discussion following the statement of their theorem 2), where again one of the forms Φ_1, Φ_2 is required to be diagonal. We remark that although Chowla & Davenport do not discuss an explicit estimate for $E(N)$, an inspection of their argument will reveal that the limit of their method would yield $E(N) \ll N^{29/30+\varepsilon}$. The strength of our estimate for $E(N)$ could certainly be improved with greater effort; the interested reader will find the salient details in §§ 3 and 4.

Our approach to theorem 1.1 is based on an application of the circle method, and may be described as an amalgam of the work of Chowla & Davenport (1961) and techniques recently developed in the theory of Waring's problem. It seems inappropriate to comment on all of the ingredients at the present stage, and we postpone a

more detailed discussion to §§ 5 and 6. We begin in § 2 by recalling the basic exponential sum estimates for binary cubic forms from Chowla & Davenport (1961), and elaborate on these ideas in the context of the difference polynomials which arise from our methods. In § 3 we study exponential sums corresponding to the binary cubic forms in mean square on suitable 'major arcs'. Our ideas here are motivated by a strategy adopted by Hooley (1986). Next, in § 4, we derive the upper bound provided in theorem 1.1. The latter turns out to be a very straightforward consequence of a suitable Weyl-type bound for exponential sums. It transpires that such bounds fail, by a factor P^{ε}, to establish the lower bound stated in theorem 1.1. In such circumstances, one can hope to prove the desired lower bound by making use of an efficient differencing process restricted to the minor arcs of the Hardy–Littlewood dissection, a process which has been successfully applied, for example by Vaughan (1986, 1989) and Vaughan & Wooley (1991, 1994), in the context of Waring's problem. We adapt this idea to handle binary cubic forms in § 5. Although differencing a polynomial in two variables is less efficient than the corresponding operation for a single variable, this 'inefficient' differencing process is nonetheless sufficient for our purposes. In order to complete our estimation of the contribution from the minor arcs, we develop an important pruning process, this making use of estimates from § 3. The proof of theorem 1.1 is concluded in § 6 with the evaluation of the major arc contribution. Finally, in § 7, we provide an outline of the proof of theorem 1.2.

An alternative strategy to establish the non-trivial solubility of the diophantine equation (1.1) may be found in work of Lewis (1957*a*). One observes that a binary cubic form, having integer coefficients and non-zero discriminant, is equivalent under a linear transformation over a quadratic field extension of $\mathbb{Q}$ to a diagonal form with coefficients in the latter field extension. Consequently, equation (1.1) is equivalent to a diagonal equation defined over a field K which arises from a succession of quadratic extensions of $\mathbb{Q}$. Modern versions of the circle method applicable to algebraic number fields should be of power sufficient to establish the solubility of a diagonal equation in eight variables, although the required conclusion appears presently to be absent from the literature. Given the existence of a non-trivial solution in K, one may use a method described by Lewis (1957*a*) to pull back, through the tower of quadratic extensions, to a non-trivial rational solution satisfying (1.1). This proposed strategy would fail, of course, to establish the fairly precise information concerning the size of $\mathcal{N}(P)$ provided by theorem 1.1.

Throughout this paper, implicit constants occurring in Vinogradov's notation $\ll$ and $\gg$ will depend at most on the coefficients of the implicit binary forms, a small positive number ε, and quantities occurring as subscripts to the latter notations, unless otherwise indicated. When $x \in \mathbb{R}$ we write $\|x\|$ for $\min_{y\in\mathbb{Z}} |x-y|$. We write $p^s \| n$ when $p^s | n$ and $p^{s+1} \nmid n$. Also, we use vector notation for brevity. Thus, for example, $(\Phi_1, \dots, \Phi_4)$ will be abbreviated simply to $\boldsymbol{\Phi}$. In an effort to simplify our exposition, we adopt the convention that whenever ε appears in a statement, we are implicitly asserting that the statement holds for each $\varepsilon > 0$. Note that the 'value' of ε may consequently change from statement to statement.

2. Exponential sums involving binary cubic forms

We pave the way for the technical aspects of our argument, described in §§ 5–7, by recording in this section a number of estimates for exponential sums over binary

cubic forms. We denote by $\Phi = \Phi(x, y)$ the binary cubic form

$$\Phi(x, y) = ax^3 + bx^2y + cxy^2 + dy^3, \tag{2.1}$$

in which a, b, c and d are fixed integers. We suppose that the discriminant of Φ, defined by

$$D = 18abcd + b^2c^2 - 4(ac^3 + db^3) - 27a^2d^2, \tag{2.2}$$

is non-zero.

Lemma 2.1. *Let $\phi(x, y)$ be any polynomial with real coefficients of degree at most 2. Let α be a real number, and suppose that there exist $r \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $(r, q) = 1$ and $|\alpha - r/q| \leqslant q^{-2}$. Then for each positive number ε,*

$$\sum_{1 \leqslant x \leqslant P} \sum_{1 \leqslant y \leqslant P} e(\alpha\Phi(x, y) + \phi(x, y)) \ll P^{2+\varepsilon}(q^{-1} + P^{-1} + qP^{-3})^{1/2}. \tag{2.3}$$

Here, the implicit constant depends at most on Φ and ε, but not on ϕ. Further, if $1 \leqslant q \leqslant P$ and $|q\alpha - r| \leqslant P^{-2}$, then

$$\sum_{1 \leqslant x \leqslant P} \sum_{1 \leqslant y \leqslant P} e(\alpha\Phi(x, y) + \phi(x, y)) \ll P^{2+\varepsilon}(q + P^3|q\alpha - r|)^{-1/2}. \tag{2.4}$$

Proof. The inequality (2.3) is immediate from theorem 1 of Chowla & Davenport (1961), and the associated conclusion (2.4) is readily derived from (2.3) via a standard argument (see, for example, Davenport & Heilbronn (1937), or exercise 2 of ch. 2 of Vaughan (1997)). ■

The observant reader will have noticed that the estimates provided by lemma 2.1 are of the same quality as those stemming from the classical inequality of Weyl for the special case in which Φ is diagonal.

We next derive a variant of lemma 2.1 relating to the difference polynomial stemming from Φ, defined for each natural number m by

$$\Psi_m(x, y; h, k) = m^{-1}(\Phi(x + hm, y + km) - \Phi(x, y)). \tag{2.5}$$

Note that Ψ_m has integral coefficients, and is quadratic in x and y. Let P be a large real number, and let H be a real number with $1 \leqslant H \leqslant Pm^{-1}$. For each h and k with $|h|, |k| \leqslant H$, let $I(h)$ and $J(k)$ be subintervals of $[-P, P]$. Define next the exponential sum $F_m(\alpha) = F_m(\alpha; \boldsymbol{I}, \boldsymbol{J})$ associated with Ψ_m by

$$F_m(\alpha) = \sum_{\substack{0 \leqslant |h|, |k| \leqslant H \\ (h,k) \neq (0,0)}} \sum_{\substack{x \in I(h) \\ y \in J(k)}} e(\alpha\Psi_m(x, y; h, k)). \tag{2.6}$$

Lemma 2.2. *Suppose that $-3D$ is not a square. Let α be a real number, and suppose that there exist $r \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $(r, q) = 1$ and $|\alpha - r/q| \leqslant q^{-2}$. Then, uniformly in $\boldsymbol{I}$ and $\boldsymbol{J}$, for each positive number ε one has*

$$F_m(\alpha; \boldsymbol{I}, \boldsymbol{J}) \ll P^{2+\varepsilon}H^2(q^{-1} + P^{-1} + q(HP^2)^{-1}). \tag{2.7}$$

When $1 \leqslant q \leqslant P$ and $|q\alpha - r| \leqslant (HP)^{-1}$, moreover, one has

$$F_m(\alpha; \boldsymbol{I}, \boldsymbol{J}) \ll P^{2+\varepsilon}H^2(q + HP^2|q\alpha - r|)^{-1}. \tag{2.8}$$

Proof. We follow the path laid down by Chowla & Davenport (1961). By applying Cauchy's inequality to (2.6), one obtains

$$|F_m(\alpha)|^2 \leqslant H^2 \sum_{\substack{0\leqslant|h|,|k|\leqslant H\\(h,k)\neq(0,0)}} \sum_{\substack{x_1\in I(h)\\y_1\in J(k)}} \sum_{\substack{x_2\in I(h)\\y_2\in J(k)}} e(\alpha(\Psi_m(x_1,y_1;h,k)-\Psi_m(x_2,y_2;h,k))).$$

Thus, on isolating the diagonal contribution, a modest calculation reveals that

$$|F_m(\alpha)|^2 \leqslant H^2(P^2+\Upsilon), \tag{2.9}$$

where

$$\Upsilon = \sum_{\substack{0\leqslant|h|,|k|\leqslant H\\(h,k)\neq(0,0)}} \sum_{\substack{0\leqslant|u|,|v|\leqslant 2P\\(u,v)\neq(0,0)}} \min\{P,\|2\alpha B_1\|^{-1}\}\min\{P,\|2\alpha B_2\|^{-1}\}, \tag{2.10}$$

with

$$B_1 = B_1(u,v;h,k) = 3ahu + b(ku+hv) + ckv,$$

and

$$B_2 = B_2(u,v;h,k) = bhu + c(ku+hv) + 3dkv.$$

When m_1 and m_2 are integers, let $\mathcal{B}(m_1,m_2)$ denote the number of solutions of the system of equations

$$B_i(u,v;h,k) = m_i \quad (i=1,2),$$

with $|h|,|k| \leqslant H$ and $|u|,|v| \leqslant 2P$, and subject to the conditions $(h,k)\neq(0,0)$ and $(u,v)\neq(0,0)$. Then by lemma 2 of Chowla & Davenport (1961), one has $\mathcal{B}(m_1,m_2) \ll (HP)^{\varepsilon}$, whence by (2.10),

$$\Upsilon \ll (HP)^{\varepsilon} \sum_{1\leqslant m_1\leqslant \kappa HP} \sum_{1\leqslant m_2\leqslant \kappa HP} \min\{P,\|2\alpha m_1\|^{-1}\}\min\{P,\|2\alpha m_2\|^{-1}\},$$

where here we write $\kappa = 12\max\{|a|,|b|,|c|,|d|\}$. On applying a standard estimate for such reciprocal sums (see, for example, lemma 2.2 of Vaughan (1997)), and recalling (2.9), we deduce that

$$|F_m(\alpha)|^2 \ll P^2H^2 + H^2(\log(2PHq))^2(PH+q+P^2Hq^{-1})^2,$$

and the upper bound (2.7) follows immediately. The estimate (2.8) follows from the latter bound by means of the same standard argument cited in the proof of lemma 2.1. ■

When $\mathcal{B}\subset\mathbb{R}^2$ is a rectangle with sides parallel to the axes, define $f(\alpha)=f(\alpha;P)$ by

$$f(\alpha;P) = \sum_{(x,y)\in P\mathcal{B}} e(\alpha\Phi(x,y)). \tag{2.11}$$

We require an approximation to $f(\alpha)$ of use on the major arcs in a Hardy–Littlewood dissection, and this will entail investigating the complete exponential sum

$$S(q,r) = \sum_{x=1}^{q}\sum_{y=1}^{q} e\left(\frac{r}{q}\Phi(x,y)\right), \tag{2.12}$$

and, when $\mathcal{D}$ is a convex subset of $\mathbb{R}^2$, the exponential integral

$$v(\beta;\mathcal{D}) = \int\!\!\int_{\mathcal{D}} e(\beta\Phi(\xi,\eta))\,\mathrm{d}\xi\mathrm{d}\eta. \tag{2.13}$$

Lemma 2.3. *Let $\mathcal{B} \subset \mathbb{R}^2$ be a box, and let δ be a positive number. Suppose that α is a real number, and that $r \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $(r,q) = 1$, $q \leqslant P^{1-\delta}$ and $|q\alpha - r| \leqslant P^{-2-\delta}$. Then*

$$f(\alpha; P) - q^{-2}S(q,r)v(\beta; P\mathcal{B}) \ll P^{1+\varepsilon}q^{1/2}.$$

Proof. This is immediate from lemmata 10 and 11 of Chowla & Davenport (1961). ■

In our subsequent investigations concerning the contribution of the major arcs in the Hardy–Littlewood dissection, we will require estimates for the generating functions defined in (2.12) and (2.13). In this context, we note that Chowla & Davenport (1961) estimated $S(q,r)$ through an appeal to lemma 2.1 (theorem 1 of their paper), this yielding the simple bound $S(q,r) = O(q^{3/2+\varepsilon})$. Such a manoeuvre enables the latter authors to avoid a detailed study of $S(q,r)$, but is possible only because of the presence of a diagonal form. We therefore continue with a more careful examination of the aforementioned generating functions.

As is familiar with complete exponential sums of arithmetic type, the sum $S(q,r)$ has a quasi-multiplicative property. Thus, whenever $(q_1q_2, r) = (q_1, q_2) = 1$, one has

$$S(q_1q_2, r) = S(q_1, rq_2^2)S(q_2, rq_1^2), \tag{2.14}$$

and so it suffices to estimate $S(q,r)$ when q is a prime power.

Lemma 2.4. *When $r \in \mathbb{Z}$ and p is a prime number with $(3Dr, p) = 1$, one has $|S(p,r)| \leqslant 9p$.*

Proof. Our strategy is to reduce $S(p,r)$ to the one-dimensional sum

$$T(p,b) = \sum_{x=1}^{p} e(bx^3/p).$$

When $(x,p) = 1$, write $\bar{x}$ for the multiplicative inverse of x modulo p, and note that $\Phi(x,y) \equiv x^3\Phi(1, y\bar{x}) \pmod p$. Then on isolating the contribution from the terms with $x = p$, and substituting $z = y\bar{x}$ for the remaining terms, one obtains

$$S(p,r) = \sum_{z=1}^{p} T(p, r\Phi(1,z)) + E_p, \tag{2.15}$$

where $|E_p| \leqslant 2p$.

Suppose first that $p \equiv 2 \pmod 3$. Then $T(p,b)$ is zero unless $p|b$, in which case $T(p,b) = p$. Thus (2.15) implies that

$$S(p,r) = p\rho(p) + E_p,$$

where $\rho(p)$ denotes the number of solutions of the congruence $\Phi(1,z) \equiv 0 \pmod p$. It may be verified that whenever $p \nmid D$, one has $\rho(p) \leqslant 3$, and hence in this case the desired conclusion follows at once.

Suppose next that $p \equiv 1 \pmod 3$. Let χ denote a non-principal character modulo p for which χ^3 is principal, and denote by $\tau(\chi)$ the associated Gauss sum. Then by lemma 4.3 of Vaughan (1997), whenever $p \nmid b$,

$$T(p,b) = \chi(b)\tau(\bar{\chi}) + \bar{\chi}(b)\tau(\chi).$$

Thus, adopting a similar argument to that of the previous paragraph, it now follows from (2.15) that

$$S(p,r)=p\rho(p)+\sum_{\substack{z=1\\ p\nmid\Phi(1,z)}}^{p}(\chi(r\Phi(1,z))\tau(\bar{\chi})+\bar{\chi}(r\Phi(1,z))\tau(\chi))+E_p. \tag{2.16}$$

However, one has $|\tau(\chi)|=\sqrt{p}$, and an estimate due to Weil (see, for example, corollary 2C′ of ch. II of Schmidt (1976)) shows that

$$\Big|\sum_{\substack{z=1\\ p\nmid r\Phi(1,z)}}^{p}\chi(r\Phi(1,z))\Big|\leqslant 2\sqrt{p}.$$

The proof of the lemma is therefore completed on inserting the latter estimate into (2.16). ■

Lemma 2.5. *When $r\in\mathbb{Z}$ and p is a prime number with $(6Dr,p)=1$, one has $S(p^2,r)=p^2$, and when $\nu\geqslant 3$ one has $S(p^\nu,r)\ll p^{4\nu/3}$.*

Proof. Suppose that r and p satisfy the hypotheses of the statement of the lemma, so that in particular one has $p>3$. We start by showing that there is no loss of generality in supposing that $p\nmid a$. First, if $p\nmid d$, then one can simply interchange the roles of x and y in $\Phi(x,y)$ to ensure instead that $p\nmid a$. Meanwhile, if $p|a$ and $p|d$, then in view of the hypothesis that $p\nmid D$ it follows from (2.2) that $p\nmid bc$. Since $p>2$ there is an integer e with $p\nmid e(b+e)$, and since $p|a$ and $p|d$ it follows from (2.1) that

$$\Phi(cx,cy+ex)\equiv b(cx)^2(cy+ex)+c(cx)(cy+ex)^2 \pmod p.$$

The coefficient of x^3 in the latter expression is $c^2e(b+e)$, which is not divisible by p, and consequently this change of variables justifies our assumption that $p\nmid a$.

Now write

$$A=9ac-3b^2,\quad B=27a^2d+2b^3-9abc,$$

and consider the form

$$\Upsilon(x,y)=x^3+Axy^2+By^3$$

which is linked with Φ via the simple identity

$$27a^2\Phi(x,y)=\Upsilon(3ax+by,y). \tag{2.17}$$

By (2.2), the discriminant of the form $\Upsilon(x,y)$ is $D_1=-(4A^3+27B^2)$, and thus it is readily confirmed that whenever $p|D_1$ and $p\nmid a$, then one has $p|3D$, contrary to our initial assumptions. Consequently, we have $(D_1,p)=1$.

Equipped with the above simplifications, we next consider the exponential sum $S(p^\nu,r)$ with $\nu\geqslant 2$ and $p\nmid r$. Substitute $(z,w)=(3ax+by,y)$. Since $p\nmid 3a$, this change of variables is non-singular modulo p. Thus, on taking s to be an integer with $27a^2s\equiv r \pmod{p^\nu}$, we deduce from (2.17) that

$$S(p^\nu,r)=\sum_{z=1}^{p^\nu}\sum_{w=1}^{p^\nu}e\left(\frac{s}{p^\nu}\Upsilon(z,w)\right). \tag{2.18}$$

Write $z = u + hp^{\nu-1}$ with $1 \leqslant u \leqslant p^{\nu-1}$ and $1 \leqslant h \leqslant p$, and make a similar substitution for w in (2.18). We obtain

$$S(p^\nu, r) = \sum_{u=1}^{p^{\nu-1}} \sum_{v=1}^{p^{\nu-1}} \sum_{h=1}^{p} \sum_{k=1}^{p} e\left(\frac{s}{p^\nu}\Upsilon(u,v)\right) e\left(\frac{s}{p}(h\Upsilon_u(u,v) + k\Upsilon_v(u,v))\right), \tag{2.19}$$

where

$$\Upsilon_u(u,v) = 3u^2 + Av^2 \quad \text{and} \quad \Upsilon_v(u,v) = 2Auv + 3Bv^2. \tag{2.20}$$

The double sum over h and k in (2.19) vanishes unless u and v satisfy the congruences

$$\Upsilon_u(u,v) \equiv \Upsilon_v(u,v) \equiv 0 \pmod p.$$

But in the latter circumstances it follows from (2.20) that

$$0 \equiv 4A^2v^2(3u^2 + Av^2) \equiv v^4(27B^2 + 4A^3) \equiv -D_1v^4 \pmod p,$$

whence the condition $(D_1, p) = 1$ implies that $p|v$. On recalling (2.20) and noting that $p > 3$, therefore, the congruence $\Upsilon_u(u,v) \equiv 0 \pmod p$ implies that necessarily $p|u$. On substituting these conditions into (2.19), we conclude that

$$S(p^\nu, r) = p^2 \sum_{\substack{u=1 \\ u\equiv 0 \pmod p}}^{p^{\nu-1}} \sum_{\substack{v=1 \\ v\equiv 0 \pmod p}}^{p^{\nu-1}} e\left(\frac{s}{p^\nu}\Upsilon(u,v)\right). \tag{2.21}$$

When $\nu = 2$ the formula (2.21) yields

$$S(p^2, r) = p^2, \tag{2.22}$$

providing the first claim of the lemma. Meanwhile, when $\nu \geqslant 3$, one obtains from (2.21) the relation

$$S(p^\nu, r) = p^2 \sum_{u=1}^{p^{\nu-2}} \sum_{v=1}^{p^{\nu-2}} e\left(\frac{s}{p^{\nu-3}}\Upsilon(u,v)\right) = p^4 \sum_{x=1}^{p^{\nu-3}} \sum_{y=1}^{p^{\nu-3}} e\left(\frac{s}{p^{\nu-3}}\Upsilon(x,y)\right).$$

On reversing our initial change of variables we deduce the recursion formula

$$S(p^\nu, r) = p^4 S(p^{\nu-3}, r),$$

and thus the second claim of the lemma follows on making use of lemma 2.4 and (2.22). ∎

When $(p, r) = 1$, we now have at our disposal efficient estimates for $S(p^\nu, r)$ provided that $p \nmid 6D$. When $p|6D$, meanwhile, a simple bound results from applying lemma 2.1 with $P = q = p^\nu$. For easy reference we summarize these results in the following lemma.

Lemma 2.6. *Suppose that $r \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $(q, r) = 1$. Let $q = q_0q_1q_2$, where q_0, q_1, q_2 are pairwise coprime, and where q_1 is cube-free, q_2 is cube-full, $(q_1q_2, 6D) = 1$, and whenever $p|q_0$ one has $p|6D$. Then*

$$S(q, r) \ll q^{2+\varepsilon} q_0^{-1/2} q_1^{-1} q_2^{-2/3}.$$

Proof. We make use of the multiplicative property (2.14) combined with the conclusions of lemmata 2.4 and 2.5, and the above observation. ∎

We now turn our attention to the exponential integral (2.13).

Lemma 2.7. *Let $\mathcal{B} \subset \mathbb{R}^2$ be a fixed rectangle with sides parallel to the axes. Then*

$$v(\beta; P\mathcal{B}) \ll P^2(1 + P^3|\beta|)^{-2/3}.$$

Proof. In advance of the main body of our argument, we establish the auxiliary estimate

$$\int_0^A \xi e(\gamma\xi^3)\,\mathrm{d}\xi \ll \min\{A^2, |\gamma|^{-2/3}\}, \tag{2.23}$$

valid for $A > 0$ and $\gamma \in \mathbb{R}$. In order to establish this bound, we note that the left-hand side of (2.23) is $O(A^2)$, by a trivial estimate, and hence we may assume that $|\gamma| > A^{-3}$. In such circumstances

$$\int_0^A \xi e(\gamma\xi^3)\,\mathrm{d}\xi = \int_{|\gamma|^{-1/3}}^A \xi e(\gamma\xi^3)\,\mathrm{d}\xi + O(|\gamma|^{-2/3}), \tag{2.24}$$

and thus the change of variable $t = \xi^3$, followed by a partial integration, shows that the integral on the right-hand side of (2.24) is $O(|\gamma|^{-2/3})$, as required.

Now we launch into the proof of the lemma proper, noting that by a change of variables in (2.13), it suffices to establish the estimate

$$v(\beta; \mathcal{B}) \ll \min\{1, |\beta|^{-2/3}\}. \tag{2.25}$$

We establish the latter estimate by reducing the integral (2.13) to one-dimensional exponential integrals, and then bring in the estimate (2.23). Note that without loss of generality we may suppose β to be positive. Further, since the bound (2.25) is trivial for $\beta < 1$, it suffices to show that when $\beta \geqslant 1$ one has

$$v(\beta; \mathcal{B}) \ll \beta^{-2/3}. \tag{2.26}$$

Since the characteristic function of any rectangle with sides parallel to the axes is a linear combination of the characteristic functions of at most four such rectangles with the origin as a common corner, we may assume that the origin is located at one of the corners of $\mathcal{B}$. It is therefore enough to consider rectangles of the shape

$$\mathcal{B} = [0, \xi_0] \times [0, \eta_0],$$

where ξ_0 and η_0 are fixed positive constants. Write $\kappa = \eta_0/\xi_0$, and dissect $\mathcal{B}$ into the disjoint union of the triangular regions

$$\mathcal{B}_1 = \{(\xi, \eta) \in \mathcal{B} : \eta \leqslant \kappa\xi\} \quad \text{and} \quad \mathcal{B}_2 = \{(\xi, \eta) \in \mathcal{B} : \eta > \kappa\xi\}.$$

Then by (2.13),

$$v(\beta; \mathcal{B}) = v(\beta; \mathcal{B}_1) + v(\beta; \mathcal{B}_2). \tag{2.27}$$

In order to estimate $v(\beta; \mathcal{B}_1)$, we first make a change of variables. Thus, on writing $\phi(t)$ for $\Phi(1, t)$ we have

$$|v(\beta; \mathcal{B}_1)| = \left| \int_0^{\xi_0} \int_0^{\kappa\xi} e(\beta\xi^3\phi(\eta/\xi))\,\mathrm{d}\eta\mathrm{d}\xi \right| = \left| \int_0^{\kappa} \int_0^{\xi_0} \xi e(\beta\xi^3\phi(t))\,\mathrm{d}\xi\mathrm{d}t \right|,$$

and so it follows from (2.23) that

$$v(\beta; \mathcal{B}_1) \ll \int_0^{\kappa} \min\{1, (\beta|\phi(t)|)^{-2/3}\}\,\mathrm{d}t. \tag{2.28}$$

Since D is non-zero, the polynomial $\phi(t)$ is either quadratic or cubic, and thus has either two or three distinct zeros in the complex plane, each zero being simple. We define a parameter C according to the number of real zeros of $\phi'(t)$ as follows. If $\phi'(t)$ has no real zero, we put $C = 1$. If $\phi'(t)$ has one or more real zeros, then we take

$$C = \tfrac{1}{2} \min\{|\phi(\tau)| : \phi'(\tau) = 0\}.$$

Making use of the simplicity of the zeros in the latter definition, one has in either case that $C > 0$. We next dissect the interval $[0, \kappa]$ into the subsets

$$\begin{aligned} \mathcal{T}_1 &= \{t \in [0,\kappa] : |\phi(t)| < \beta^{-2/3}\}, \\ \mathcal{T}_2 &= \{t \in [0,\kappa] : \beta^{-2/3} \leqslant |\phi(t)| \leqslant C\}, \\ \mathcal{T}_3 &= \{t \in [0,\kappa] : |\phi(t)| > C\}, \end{aligned}$$

and aim to establish that for $1 \leqslant j \leqslant 3$,

$$\int_{\mathcal{T}_j} \min\{1, (\beta|\phi(t)|)^{-2/3}\}\, \mathrm{d}t \ll |\beta|^{-2/3}. \tag{2.29}$$

Since ϕ has all of its zeros simple, the measure of $\mathcal{T}_1$ is $O(\beta^{-2/3})$, and hence (2.29) is immediate when $j = 1$. Next we consider the case $j = 2$. Here we note that if $\phi'(\tau) = 0$, then $\tau \notin \mathcal{T}_2$, whence $\inf_{t\in\mathcal{T}_2} |\phi'(t)| > 0$. Moreover, since ϕ is a quadratic or cubic polynomial, the set $\mathcal{T}_2$ is a union of at most six intervals, on each of which $\phi(t)$ is monotone. If I denotes any such interval, the change of variable $u = \phi(t)$ yields

$$\int_I \min\{1, (\beta|\phi(t)|)^{-2/3}\}\, \mathrm{d}t \ll \beta^{-2/3} \int_{\beta^{-2/3}}^{C} u^{-2/3}\, \mathrm{d}u \ll \beta^{-2/3},$$

whence (2.29) follows in the case $j = 2$. When $j = 3$ the estimate (2.29) is trivial (though here, one should note, the implicit constant depends on κ, which in turn depends at most on the coefficients of Φ). On combining (2.28) and (2.29), therefore, we may conclude that

$$v(\beta; \mathcal{B}_1) \ll \beta^{-2/3}. \tag{2.30}$$

Finally, the bound $v(\beta; \mathcal{B}_2) \ll \beta^{-2/3}$ follows via the same argument as that used to bound $v(\beta; \mathcal{B}_1)$, on interchanging the roles of ξ and η. On recalling (2.27) and (2.30), therefore, we at last deduce the upper bound (2.25), and thus the proof of the lemma is complete. ■

3. Weighted exponential sums and a mean value estimate

The exponential sum $f(\alpha)$ defined in (2.11) can be approximated, when α is close to a rational point a/q, by the expected main term. However, the bound for the error arising in this approximation which follows from lemma 2.3 is unusually large, owing to the presence of the factor P. In this section we aim to show, roughly speaking, that this error is much smaller in mean square. Our argument will make use of a two-dimensional version of the Poisson summation formula, the application of which will be much facilitated by considering a weighted variant of the exponential sum discussed in the previous section.

In this section we continue to suppose that $\Phi(x,y)$ is defined by (2.1), and that the discriminant of Φ is non-zero. The Hessian of $\Phi(x,y)$ is the quadratic form

$$\mathcal{H}(x,y) = \det \begin{pmatrix} 6ax+2by & 2bx+2cy \\ 2bx+2cy & 2cx+6dy \end{pmatrix}.$$

The real locus of zeros of the polynomial $\mathcal{H}(x,y)$ either consists of the single point $(0,0)$, or else is the union of two lines through $(0,0)$. Let $(\xi_0,\eta_0) \in \mathbb{R}^2$ be any point with the property that $\mathcal{H}(\xi,\eta)$ is non-zero whenever $|\xi-\xi_0| \leqslant 1$ and $|\eta-\eta_0| \leqslant 1$. We will describe a point of the latter type as *admissible* for Φ. In view of the preceding comments, one has that every point on the real plane is admissible for Φ, except possibly for the union of the unit width neighbourhoods around two lines through the origin. Now let

$$\gamma(t) = \begin{cases} \exp\left(\dfrac{2}{t^2-1}\right), & \text{when } |t|<1, \\ 0, & \text{otherwise,} \end{cases} \tag{3.1}$$

and define, for $Q \geqslant 1$, the two-dimensional weight function $\Gamma(\xi,\eta) = \Gamma_Q(\xi,\eta)$ by

$$\Gamma_Q(\xi,\eta) = \gamma\left(\frac{\xi}{Q}-\xi_0\right)\gamma\left(\frac{\eta}{Q}-\eta_0\right). \tag{3.2}$$

The weighted exponential sum which forms the basis for our analysis is $g(\alpha) = g(\alpha;Q)$, which we define by

$$g(\alpha;Q) = \sum_{(x,y)\in\mathbb{Z}^2} \Gamma_Q(x,y)e(\alpha\Phi(x,y)).$$

On applying the Poisson summation formula, one deduces that whenever $r \in \mathbb{Z}$, $q \in \mathbb{N}$, $(r,q)=1$ and $\beta \in \mathbb{R}$, one has

$$g\left(\frac{r}{q}+\beta\right) = q^{-2} \sum_{(h,k)\in\mathbb{Z}^2} S(q,r;h,k)W\left(\beta;\frac{h}{q},\frac{k}{q}\right), \tag{3.3}$$

where

$$S(q,r;h,k) = \sum_{x=1}^{q}\sum_{y=1}^{q} e\left(\frac{r\Phi(x,y)+hx+ky}{q}\right) \tag{3.4}$$

and

$$W(\beta;u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Gamma_Q(\xi,\eta)e(\beta\Phi(\xi,\eta)-u\xi-v\eta)\,\mathrm{d}\xi\mathrm{d}\eta.$$

Note that $S(q,r;0,0) = S(q,r)$. For the sake of concision, write

$$w(\beta) = W(\beta;0,0). \tag{3.5}$$

We next define a Hardy–Littlewood dissection. Take R to be a parameter with $1 \leqslant R \leqslant \frac{1}{2}Q^{3/2}$, and when $r \in \mathbb{Z}$ and $q \in \mathbb{N}$, write

$$\mathfrak{N}(q,r) = \{\alpha \in [0,1) : |q\alpha - r| \leqslant RQ^{-3}\}.$$

We take $\mathfrak{N}(R)$ to be the union of the intervals $\mathfrak{N}(q,r)$ with $0 \leqslant r \leqslant q \leqslant R$ and $(r,q)=1$. Note that the intervals occurring in the latter union are disjoint. Finally, when $\alpha \in \mathfrak{N}(q,r) \subseteq \mathfrak{N}(R)$, define

$$E(\alpha) = g(\alpha) - q^{-2}S(q,r)w(\alpha-r/q). \tag{3.6}$$

We are now in a position to state the main result of this section.

Lemma 3.1. *Let $1 \leqslant R \leqslant \frac{1}{2}Q^{3/2}$, and suppose that (ξ_0, η_0) is an admissible point. Then*

$$\int_{\mathfrak{N}(R)} |E(\alpha)|^2 \,\mathrm{d}\alpha \ll R^{9/2}Q^{\varepsilon-3} + R^{1/2}Q^{1+\varepsilon}.$$

The interested reader may care to compare lemma 3.1 with theorem 2 of Brüdern (1991), where a stronger bound is established in the special case in which $\Phi(x, y) = x^3 + y^3$. The basic idea of the proof is modelled along the lines of Hooley (1986), but in the present situation we avoid reference to deeper results from algebraic geometry. With additional effort one should be able to improve substantially on the elementary estimate provided here.

In order to establish lemma 3.1, we require some auxiliary estimates which we describe below. When h_i and k_i are integers $(i = 1, 2)$, define

$$N(q; \boldsymbol{h}, \boldsymbol{k}) = \sum_{\substack{r=1 \\ (r,q)=1}}^{q} S(q, r; h_1, k_1)\bar{S}(q, r; h_2, k_2). \tag{3.7}$$

Lemma 3.2. *For fixed integers h_i and k_i $(i = 1, 2)$, one has that $N(q; \boldsymbol{h}, \boldsymbol{k})$ is a multiplicative function of q. Moreover, if $q = q_1q_2$ with q_1 square-free, q_2 square-full and $(q_1, q_2) = 1$, then*

$$N(q; \boldsymbol{h}, \boldsymbol{k}) \ll q^{\varepsilon} q_1^3 q_2^4 (q_1, h_1, k_1, h_2, k_2).$$

Proof. That $N(q; \boldsymbol{h}, \boldsymbol{k})$ is a multiplicative function of q follows immediately via well-known methods, and we take the liberty of omitting the details here (but see, for example, the argument of the proof of lemma 2.11 of Vaughan (1997)). Moreover, on applying lemma 2.1 to (3.4), it follows that when $(r, q) = 1$ we have

$$S(q, r; h, k) \ll q^{3/2+\varepsilon},$$

whence by (3.7),

$$N(q; \boldsymbol{h}, \boldsymbol{k}) \ll q^{4+\varepsilon}. \tag{3.8}$$

Observe that the estimate (3.8) suffices to establish the lemma when q is square-full, and also when q is a prime number dividing all of the h_i and k_i $(i = 1, 2)$. In view of the multiplicative property of $N(q; \boldsymbol{h}, \boldsymbol{k})$, therefore, the conclusion of the lemma will follow provided we show that when p is a prime number with $(p, h_1, k_1, h_2, k_2) = 1$, then one has

$$N(p; \boldsymbol{h}, \boldsymbol{k}) \ll p^3. \tag{3.9}$$

We establish (3.9) by following an argument used in work of Hooley (see the proof of lemma 6 of Hooley (1986)). Observe that by a change of variable,

$$\begin{aligned} N(p; \boldsymbol{h}, \boldsymbol{k}) &= \frac{1}{p-1} \sum_{l=1}^{p-1} \sum_{r=1}^{p-1} S(p, rl^3; h_1l, k_1l)\bar{S}(p, rl^3; h_2l, k_2l) \\ &= \frac{1}{p-1} \sum_{l=1}^{p-1} \Bigg(\sum_{r=1}^{p} S(p, r; h_1l, k_1l)\bar{S}(p, r; h_2l, k_2l) - \Sigma \Bigg), \end{aligned} \tag{3.10}$$

where

$$\Sigma = S(p, 0; h_1l, k_1l)\bar{S}(p, 0; h_2l, k_2l).$$

But by assumption, at least one of the h_i or k_i is not divisible by p, so that Σ necessarily vanishes. Thus we deduce from (3.10) that

$$N(p;\boldsymbol{h},\boldsymbol{k}) = \frac{p}{p-1}\sum_{l=1}^{p-1}\sum_{\boldsymbol{x},\boldsymbol{y}} e\left(\frac{l}{p}(h_1x_1+k_1y_1-h_2x_2-k_2y_2)\right), \tag{3.11}$$

where the summation is over x_i and y_i with

$$1 \leqslant x_i, y_i \leqslant p \quad (i=1,2) \tag{3.12}$$

subject to

$$\Phi(x_1,y_1) \equiv \Phi(x_2,y_2) \pmod p. \tag{3.13}$$

Consequently, on writing $\nu(p)$ for the number of solutions $\boldsymbol{x},\boldsymbol{y}$ of (3.13) satisfying (3.12), and $\nu(p;\boldsymbol{h},\boldsymbol{k})$ for the corresponding number of solutions subject to the additional condition

$$h_1x_1+k_1y_1 \equiv h_2x_2+k_2y_2 \pmod p,$$

we conclude from (3.11) that

$$N(p;\boldsymbol{h},\boldsymbol{k}) = \frac{p}{p-1}(p\nu(p;\boldsymbol{h},\boldsymbol{k})-\nu(p)). \tag{3.14}$$

But, again because at least one of the h_i or k_i is not divisible by p, one finds without difficulty that $\nu(p;\boldsymbol{h},\boldsymbol{k}) \ll p^2$. Thus, the estimate $\nu(p) \ll p^3$ being trivial, we conclude from (3.14) that the estimate (3.9) does indeed hold. This completes the proof of the lemma. ■

Bounds for $W(\beta;u,v)$ are available from the literature. We summarize some special cases of lemma 7 of Hooley (1988) in the following lemma.

Lemma 3.3. *Let β, u and v be real numbers, and suppose that (ξ_0,η_0) is admissible for Φ. Then whenever $|\beta| \leqslant Q^{-3/2}$, one has*

$$W(\beta;u,v) \ll Q^2(1+Q^3(\log Q)^{-2}|\beta|)^{-1}.$$

Further, there is a positive constant C such that whenever $|u|+|v| \geqslant C^{-1}|\beta|Q^2$, one has

$$W(\beta;u,v) \ll Q^2\exp(-C\sqrt{Q(|u|+|v|)}).$$

Having provided the necessary prerequisites, the stage is now set for our proof of lemma 3.1.

The proof of lemma 3.1. We begin with a rearrangement. When $\alpha \in \mathfrak{N}(R)$, it follows from (3.3) and (3.6) that

$$E(\alpha) = \sum_{(h,k)\in\mathbb{Z}^2\setminus\{(0,0)\}} q^{-2}S(q,r;h,k)W\left(\beta;\frac{h}{q},\frac{k}{q}\right).$$

Thus

$$\int_{\mathfrak{N}(R)} |E(\alpha)|^2\,\mathrm{d}\alpha = \sum_{1\leqslant q\leqslant R}\sum_{\substack{(h_1,k_1)\in\mathbb{Z}^2\setminus\{(0,0)\}\\(h_2,k_2)\in\mathbb{Z}^2\setminus\{(0,0)\}}} q^{-4}N(q;\boldsymbol{h},\boldsymbol{k})\mathcal{I}(q;\boldsymbol{h},\boldsymbol{k}), \tag{3.15}$$

where

$$\mathcal{I}(q;\boldsymbol{h},\boldsymbol{k}) = \int_{-R/(qQ^3)}^{R/(qQ^3)} W\left(\beta;\frac{h_1}{q},\frac{k_1}{q}\right)\overline{W}\left(\beta;\frac{h_2}{q},\frac{k_2}{q}\right)\,\mathrm{d}\beta. \tag{3.16}$$

Recall our conventions concerning q, q_1 and q_2 described in the statement of lemma 3.2. Then on noting that for square-free q_1 one has

$$(q_1,h_1,k_1,h_2,k_2) \leqslant (q_1,h_1,k_1)^{1/2}(q_1,h_2,k_2)^{1/2},$$

we deduce from lemma 3.2 together with (3.15) and (3.16) that

$$\int_{\mathfrak{N}(R)} |E(\alpha)|^2\,\mathrm{d}\alpha \ll R^{\varepsilon}\sum_{1\leqslant q\leqslant R} q_1^{-1}\int_{-R/(qQ^3)}^{R/(qQ^3)} V(\beta;q)^2\,\mathrm{d}\beta, \tag{3.17}$$

where

$$V(\beta;q) = \sum_{(h,k)\in\mathbb{Z}^2\setminus\{(0,0)\}}\left|W\left(\beta;\frac{h}{q},\frac{k}{q}\right)\right|(q_1,h,k)^{1/2}. \tag{3.18}$$

Before proceeding further we dock the tails from the summation over h and k occurring in (3.18). To this end, when $\beta\in\mathbb{R}$ define $H = H(\beta)$ by

$$H = \begin{cases} \max\{1, q/Q\}, & \text{when } |\beta| \leqslant Q^{-3}, \\ \max\{1, qQ^2|\beta|\}, & \text{when } Q^{-3} < |\beta| < q^{-1}Q^{-3}R. \end{cases} \tag{3.19}$$

Let δ be a fixed positive number to be chosen later. Then by lemma 3.3, there is a positive constant C such that whenever $|h|+|k| > HQ^{\delta}$ one has

$$\left|W\left(\beta;\frac{h}{q},\frac{k}{q}\right)\right|(q_1,h,k)^{1/2} \ll q_1^{1/2}Q^2\exp(-C\sqrt{Qq^{-1}(|h|+|k|)}). \tag{3.20}$$

But

$$\begin{aligned}\sum_{|h|+|k|>HQ^{\delta}} \exp(-C\sqrt{Qq^{-1}(|h|+|k|)}) &\ll \sum_{l>HQ^{\delta}} l\exp(-C(Qq^{-1}l)^{1/2}) \\ &\ll \exp(-Q^{\delta/3}),\end{aligned}$$

whence by combining lemma 3.3 with (3.18) and (3.20), we deduce that whenever $|\beta| \leqslant Q^{-3/2}$ one has

$$\begin{aligned} V(\beta;q) &= \sum_{0<|h|+|k|\leqslant HQ^{\delta}}\left|W\left(\beta;\frac{h}{q},\frac{k}{q}\right)\right|(q_1,h,k)^{1/2} + O(\exp(-Q^{\delta/4})) \\ &\ll q_1^{\varepsilon}H^2Q^{2+2\delta+\varepsilon}(1+Q^3|\beta|)^{-1} + O(\exp(-Q^{\delta/4})). \end{aligned} \tag{3.21}$$

We next substitute (3.21) into (3.17), obtaining for $1 \leqslant R \leqslant Q^{3/2}$ the estimate

$$\int_{\mathfrak{N}(R)} |E(\alpha)|^2\,\mathrm{d}\alpha \ll 1 + Q^{\varepsilon}\sum_{1\leqslant q\leqslant R} q_1^{-1}\int_{-R/(qQ^3)}^{R/(qQ^3)} H^4Q^{4+4\delta}(1+Q^3|\beta|)^{-2}\,\mathrm{d}\beta. \tag{3.22}$$

But on recalling (3.19), one readily confirms that when $1 \leqslant q \leqslant R$ one has

$$\int_{-R/(qQ^3)}^{R/(qQ^3)} H^4Q^4(1+Q^3|\beta|)^{-2}\,\mathrm{d}\beta \ll Q + qR^3Q^{-3}.$$

Further, trivially,

$$\sum_{1\leqslant q\leqslant R} q_1^{-1} \ll R^{1/2}.$$

Thus (3.22) yields

$$\int_{\mathfrak{N}(R)} |E(\alpha)|^2 \,\mathrm{d}\alpha \ll 1 + Q^{4\delta+\varepsilon}(R^{1/2}Q + R^{9/2}Q^{-3}),$$

and the proof of the lemma is completed on taking $\delta = \varepsilon$. ■

Before leaving this discussion of the behaviour of $E(\alpha)$, we extract a further simple estimate from lemma 3.3.

Lemma 3.4. *Let $\delta > 0$ and suppose that $R \leqslant Q^{1-\delta}$. Then whenever $\alpha \in \mathfrak{N}(R)$ one has $E(\alpha) \ll 1$.*

Proof. Suppose that $1 \leqslant q \leqslant Q^{1-\delta}$ and $|q\alpha - r| \leqslant Q^{-2-\delta}$. Then by lemma 3.3,

$$\begin{aligned}\sum_{(h,k)\in\mathbb{Z}^2\setminus\{(0,0)\}} \left|W\left(\beta; \frac{h}{q}, \frac{k}{q}\right)\right| &\ll Q^2 \sum_{(h,k)\in\mathbb{Z}^2\setminus\{(0,0)\}} \exp(-C\sqrt{Qq^{-1}(|h|+|k|)}) \\ &\ll Q^2 \sum_{l=1}^{\infty} l \exp(-CQ^{\delta/2}l^{1/2}) \\ &\ll \exp(-Q^{\delta/3}).\end{aligned}$$

The desired conclusion therefore follows from (3.3) and (3.6). ■

4. The proof of theorem 1.1: the upper bound

We establish the upper bound provided in theorem 1.1 by a straightforward argument based on the use of lemma 2.1. Let Φ be a binary cubic form with non-zero discriminant, and define $f(\alpha) = f(\alpha; \Phi)$ by

$$f(\alpha; \Phi) = \sum_{|x|\leqslant P} \sum_{|y|\leqslant P} e(\alpha\Phi(x, y)).$$

Write also

$$I(\Phi) = \int_0^1 |f(\alpha; \Phi)|^4 \,\mathrm{d}\alpha. \tag{4.1}$$

Then on recalling the statement of theorem 1.1, and applying Hölder's inequality, one obtains

$$\mathcal{N}(P; \boldsymbol{\Phi}) = \int_0^1 \prod_{i=1}^{4} f(\alpha; \Phi_i) \,\mathrm{d}\alpha \leqslant \left(\prod_{i=1}^{4} I(\Phi_i)\right)^{1/4}.$$

The upper bound of theorem 1.1 is therefore immediate from the bound on $I(\Phi)$ provided in the following lemma.

Lemma 4.1. *Let Φ be a binary cubic form with non-zero discriminant. Then $I(\Phi) \ll P^{5+\varepsilon}$.*

Proof. We apply the Hardy–Littlewood method. When $r \in \mathbb{Z}$ and $q \in \mathbb{N}$, write

$$\mathfrak{M}(q,r) = \{\alpha \in [0,1) : |q\alpha - r| \leqslant P^{-2}\}.$$

Take $\mathfrak{M}$ to be the union of the intervals $\mathfrak{M}(q,r)$ with $0 \leqslant r \leqslant q \leqslant P$ and $(r,q) = 1$. Note that the intervals occurring in the latter union are disjoint. Finally, write $\mathfrak{m} = [0,1) \setminus \mathfrak{M}$. We begin by noting that lemma 2.1 implies the estimate

$$\sup_{\alpha \in \mathfrak{m}} |f(\alpha; \Phi)| \ll P^{3/2+\varepsilon}.$$

Thus, in view of the upper bound

$$\int_0^1 |f(\alpha;\Phi)|^2 \,\mathrm{d}\alpha \ll P^{2+\varepsilon}, \tag{4.2}$$

which follows from lemma 5 of Chowla & Davenport (1961) (see also Hooley (1967, 1985) for important refinements of the latter), we obtain

$$\int_{\mathfrak{m}} |f(\alpha;\Phi)|^4 \,\mathrm{d}\alpha \ll (\sup_{\alpha \in \mathfrak{m}} |f(\alpha;\Phi)|)^2 \int_0^1 |f(\alpha;\Phi)|^2 \,\mathrm{d}\alpha \ll P^{5+\varepsilon}. \tag{4.3}$$

The contribution from the major arcs is readily estimated. Indeed, when $\alpha \in \mathfrak{M}(q,r) \subseteq \mathfrak{M}$, it follows from the estimate (2.4) of lemma 2.1 that

$$f(\alpha;\Phi) \ll P^{2+\varepsilon}(q + P^3|q\alpha - r|)^{-1/2},$$

and therefore,

$$\begin{aligned}\int_{\mathfrak{M}} |f(\alpha;\Phi)|^4 \,\mathrm{d}\alpha &\ll P^{8+\varepsilon} \sum_{1 \leqslant q \leqslant P} \sum_{\substack{r=1 \\ (r,q)=1}}^{q} \int_{|\beta| \leqslant (qP^2)^{-1}} (q + P^3 q|\beta|)^{-2} \,\mathrm{d}\beta \\ &\ll P^{5+\varepsilon} \sum_{1 \leqslant q \leqslant P} \sum_{r=1}^{q} q^{-2} \ll P^{5+\varepsilon}. \end{aligned} \tag{4.4}$$

The proof of the lemma is completed on combining (4.1), (4.3) and (4.4). ■

5. A minor arc estimate

The observant reader will have noticed that the minor arc estimate (4.3) exceeds the expected order of magnitude of $\mathcal{N}(P;\boldsymbol{\Phi})$ only by a factor of P^{ε}. In such circumstances, a further saving can be wrought by employing a process nowadays known as 'efficient differencing' restricted to minor arcs. This method has its roots in the work of Vaughan (1986). However, as opposed to all previous applications of this technique, we are forced to perform the differencing operation on two variables simultaneously. This operation is not quite as efficient as in more familiar applications of the method, but nonetheless stronger than making use of ordinary Weyl estimates, as in (4.3).

In order to make use of our differencing argument we require a simple lemma concerning the number of solutions of certain congruence relations.

Lemma 5.1. *Let Φ be an integral binary cubic form with non-zero discriminant D. Also, when l is an integer, write $\mathcal{A}(l)$ for the set of solutions of the congruence*

$$\Phi(x,y) \equiv l \pmod{p^3} \tag{5.1}$$

with $1 \leqslant x, y \leqslant p^3$ and $(x,y,p) = 1$. Then $\mathrm{card}(\mathcal{A}(l)) \ll p^3$, where the implicit constant depends at most on D.

Proof. By symmetry, it suffices to estimate the number of solutions of (5.1) lying in $\mathcal{A}(l)$ for which $p \nmid x$. We note also that the lemma is trivial when $p|D$, and thus we henceforth assume that $p \nmid D$. Let $\mathcal{M}(p; \Phi)$ denote the number of solutions of the congruence

$$x^3\Phi(1, z) \equiv l \pmod{p^3} \tag{5.2}$$

with $1 \leqslant x, z \leqslant p^3$ and $p \nmid x$. Then on making the substitution $y \equiv zx \pmod{p^3}$, it is apparent that the lemma will follow on showing that $\mathcal{M}(p; \Phi) \ll p^3$. In order to establish the latter estimate, we consider the set

$$\mathcal{Z} = \{1 \leqslant z \leqslant p^3 : \Phi'(1, z) \equiv 0 \pmod{p}\},$$

where $\Phi'(1, z)$ denotes the derivative of $\Phi(1, z)$, and divide into cases.

(i) *Solutions of* (5.2) *with* $z \in \mathcal{Z}$. Observe that the simultaneous congruences

$$\Phi(1, z) \equiv \Phi'(1, z) \equiv 0 \pmod{p}$$

have no solution, for otherwise $\Phi(1, z)$ would necessarily possess a double root modulo p, contradicting our assumption that $p \nmid D$. Thus we deduce that whenever $z \in \mathcal{Z}$, one has $\Phi(1, z) \not\equiv 0 \pmod{p}$. Consequently, when $p|l$ there are no solutions of (5.2). Suppose, on the other hand, that $p \nmid l$. Then for each fixed $z \in \mathcal{Z}$ there are at most three solutions in x to the congruence (5.2), whence the total number of solutions in this case is at most $3 \cdot \mathrm{card}(\mathcal{Z}) \leqslant 3p^3$.

(ii) *Solutions of* (5.2) *with* $z \notin \mathcal{Z}$. Suppose that x is an integer with $1 \leqslant x \leqslant p^3$ and $p \nmid x$. Then on writing l' for $l\bar{x}^3 \pmod{p^3}$, we find that the number of solutions of (5.2) with $z \notin \mathcal{Z}$ is at most

$$p^3 \,\mathrm{card}(\{z \notin \mathcal{Z} : \Phi(1, z) \equiv l' \pmod{p^3}\}). \tag{5.3}$$

However, there are at most three solutions of the congruence $\Phi(1, z) \equiv l' \pmod{p}$, and for each such solution with $z \notin \mathcal{Z}$ we have $p \nmid \Phi'(1, z)$. Thus a Hensel's lemma argument reveals that

$$\mathrm{card}(\{z \notin \mathcal{Z} : \Phi(1, z) \equiv l' \pmod{p^3}\}) \leqslant 3.$$

On recalling (5.3), we find that the number of solutions in this case is also at most $3p^3$.

The proof of the lemma is completed on combining the conclusions of cases (i) and (ii). ■

We now set the scene for the enunciation of a technical minor arc estimate used in the proofs of theorems 1.1 and 1.2. In addition to the idea of efficient differencing, the work described in § 3 also plays a crucial role in our argument. As usual, let Φ be an integral binary cubic form, defined as in (2.1), and having non-zero discriminant D. Let $\mathcal{B} \subset \mathbb{R}^2$ be a rectangle, and define $f(\alpha) = f(\alpha; P)$ as in (2.11). Let $\tilde{\Phi}$ denote a second integral binary cubic form with non-zero discriminant $\tilde{D}$. Let $(\tilde{\xi}, \tilde{\eta})$ be an admissible point for $\tilde{\Phi}$, and define the weight function

$$\tilde{\Gamma}_Q(\xi, \eta) = \gamma\left(\frac{\xi}{Q} - \tilde{\xi}\right)\gamma\left(\frac{\eta}{Q} - \tilde{\eta}\right). \tag{5.4}$$

Finally, denote by $g(\alpha) = g(\alpha; Q)$ the exponential sum

$$g(\alpha; Q) = \sum_{(x,y)\in\mathbb{Z}^2} \tilde{\Gamma}_Q(x, y)e(\alpha\tilde{\Phi}(x, y)). \tag{5.5}$$

Lemma 5.2. *Suppose that $-3D$ is not a square. Let p denote a prime number with $p \leqslant P^{1/10}$, and let R denote a real parameter with $R \geqslant (\frac{1}{2}p)^4$. Write $\mathfrak{m}(R)$ for the set of all $\alpha \in [0,1)$ satisfying the property that whenever $\|q\alpha\| \leqslant RP^{-3}$, then one has $q > R$. Then*

$$\int_{\mathfrak{m}(R)} |f(\alpha;P)g(\alpha p^3;P/p)|^2 \,\mathrm{d}\alpha \ll P^{5+\varepsilon}p^{-5}. \tag{5.6}$$

Observe that if the integral in (5.6) were evaluated over the interval $[0,1)$, then its expected order of magnitude would be P^5p^{-4}. Thus, whenever p is a suitably small power of P, we essentially save a factor of p on the minor arcs relative to the major arc contribution.

The proof of lemma 5.2. We start proceedings with the elimination of certain common factors amongst variables. Write $Q = P/p$, let R denote a real parameter with $R \geqslant (\frac{1}{2}p)^4$, and write $\mathfrak{m} = \mathfrak{m}(R)$. Define

$$f_p(\alpha) = \sum_{\substack{(x,y)\in P\mathcal{B}\\(x,y,p)=1}} e(\alpha\Phi(x,y)).$$

Then one has

$$f(\alpha;P) = f_p(\alpha) + f(\alpha p^3;Q). \tag{5.7}$$

Let

$$\Omega = 2 + |\tilde{\xi}| + |\tilde{\eta}| + \sup_{(\xi,\eta)\in\mathcal{B}} (|\xi| + |\eta|),$$

and write

$$\tilde{g}(\alpha) = \sum_{|x|\leqslant\Omega Q} \sum_{|y|\leqslant\Omega Q} e(\alpha\tilde{\Phi}(x,y)).$$

Then by applying Schwarz's inequality, considering the underlying diophantine equations and recalling (3.1), (5.4) and (5.5), one finds that

$$\begin{aligned}\int_0^1 |f(\alpha p^3;Q)g(\alpha p^3;Q)|^2 \,\mathrm{d}\alpha &\leqslant \left(\int_0^1 |f(\alpha p^3;Q)|^4 \,\mathrm{d}\alpha\right)^{1/2}\left(\int_0^1 |g(\alpha p^3;Q)|^4 \,\mathrm{d}\alpha\right)^{1/2}\\ &\ll \left(\int_0^1 |f(\alpha;Q)|^4 \,\mathrm{d}\alpha\right)^{1/2}\left(\int_0^1 |\tilde{g}(\alpha)|^4 \,\mathrm{d}\alpha\right)^{1/2}.\end{aligned}$$

Thus it follows from lemma 4.1 that

$$\int_0^1 |f(\alpha p^3;Q)g(\alpha p^3;Q)|^2 \,\mathrm{d}\alpha \ll Q^{5+\varepsilon}.$$

In view of (5.7) and the inequality $|u+v|^2 \leqslant 2(|u|^2 + |v|^2)$, therefore, one has

$$\int_{\mathfrak{m}} |f(\alpha;P)g(\alpha p^3;Q)|^2 \,\mathrm{d}\alpha \leqslant 2\int_{\mathfrak{m}} |f_p(\alpha)g(\alpha p^3;Q)|^2 \,\mathrm{d}\alpha + O(Q^{5+\varepsilon}). \tag{5.8}$$

Having prepared the ground, we now apply an argument first applied by Vaughan in work on Waring's problem for cubes (see lemma 10 of Vaughan (1986)). We write $R_0 = Rp^{-3}$, and let $\mathfrak{n}$ denote the set of $\alpha \in [0,1)$ satisfying the condition that

whenever q is a natural number with $\|q\alpha\| \leqslant R_0 Q^{-3}$, then one has $q > R_0$. When $k \in \mathbb{N}$, write

$$\mathfrak{n}_k = \{\alpha \in \mathbb{R} : \alpha - k \in \mathfrak{n}\}.$$

Write, further,

$$d = p^3, \quad \mathcal{M}_d = \bigcup_{k=0}^{d-1} \mathfrak{n}_k \quad \text{and} \quad \mathcal{A}_d = \{\alpha \in \mathbb{R} : \alpha d \in \mathcal{M}_d\}.$$

Then following the argument of the proof of lemma 10 of Vaughan (1986) (see also § 4 of Vaughan & Wooley (1991)), one finds that $\mathfrak{m} \subset \mathcal{A}_d$, whence it follows that

$$\int_{\mathfrak{m}} |f_p(\alpha) g(\alpha p^3; Q)|^2 \, \mathrm{d}\alpha \leqslant \int_{\mathfrak{n}} \Theta(\alpha) |g(\alpha; Q)|^2 \, \mathrm{d}\alpha, \tag{5.9}$$

where

$$\Theta(\alpha) = d^{-1} \sum_{k=0}^{d-1} \left| f_p\left(\frac{\alpha + k}{d}\right)\right|^2.$$

Moreover, by orthogonality one has

$$\Theta(\alpha) = \sum_{x,y,u,v} e(\alpha p^{-3}(\Phi(x,y) - \Phi(u,v))),$$

where the summation is over

$$(x,y) \in P\mathcal{B} \quad \text{and} \quad (u,v) \in P\mathcal{B} \tag{5.10}$$

satisfying

$$(x,y,p) = (u,v,p) = 1 \quad \text{and} \quad \Phi(x,y) \equiv \Phi(u,v) \pmod{p^3}.$$

Consequently, on recalling the notation of the statement of lemma 5.1 and applying Cauchy's inequality, one obtains

$$\begin{aligned}\Theta(\alpha) &= \sum_{l=1}^{p^3} \left| \sum_{(x_0,y_0) \in \mathcal{A}(l)} \sum_{x \equiv x_0 \, (\mathrm{mod}\, p^3)} \sum_{y \equiv y_0 \, (\mathrm{mod}\, p^3)} e(\alpha p^{-3} \Phi(x,y)) \right|^2 \\ &\leqslant \sum_{l=1}^{p^3} \mathrm{card}(\mathcal{A}(l)) \sum_{(x_0,y_0) \in \mathcal{A}(l)} \left| \sum_{x \equiv x_0 \, (\mathrm{mod}\, p^3)} \sum_{y \equiv y_0 \, (\mathrm{mod}\, p^3)} e(\alpha p^{-3} \Phi(x,y)) \right|^2,\end{aligned}$$

where the summations over x and y are subject to (5.10). We therefore deduce from lemma 5.1 that

$$\Theta(\alpha) \ll p^3 \sum_{x,y,u,v} e(\alpha p^{-3}(\Phi(x,y) - \Phi(u,v))), \tag{5.11}$$

where the summation is over x, y, u, v satisfying (5.10) with

$$(x,y,p) = 1, \quad x \equiv u \pmod{p^3} \quad \text{and} \quad y \equiv v \pmod{p^3}. \tag{5.12}$$

We now reduce the exponential sum on the right-hand side of (5.11) to related sums to which lemma 2.2 is applicable. We first remove the diagonal contribution arising from the terms with $(x,y) = (u,v)$. Next we remove the coprimality condition

$(x, y, p) = 1$, noting that the congruence conditions recorded in (5.12) imply that whenever $p|x$ and $p|y$, then one has $p|u$ and $p|v$. Thus we deduce from (5.11) that

$$\Theta(\alpha) \ll p^3(P^2 + |F(\alpha)| + |G(\alpha)|), \tag{5.13}$$

where

$$F(\alpha) = \sum_{(x,y)\in P\mathcal{B}} \sum_{\substack{(u,v)\in P\mathcal{B}\\ u\equiv x \,(\mathrm{mod}\, p^3)\\ v\equiv y \,(\mathrm{mod}\, p^3)\\ (u,v)\neq(x,y)}} e(\alpha p^{-3}(\Phi(x,y) - \Phi(u,v))) \tag{5.14}$$

and

$$G(\alpha) = \sum_{(x,y)\in Q\mathcal{B}} \sum_{\substack{(u,v)\in Q\mathcal{B}\\ u\equiv x \,(\mathrm{mod}\, p^2)\\ v\equiv y \,(\mathrm{mod}\, p^2)\\ (u,v)\neq(x,y)}} e(\alpha(\Phi(x,y) - \Phi(u,v))). \tag{5.15}$$

However, by recalling (3.1), (5.4) and (5.5), and considering the underlying diophantine equations, we may infer from (4.2) that

$$\int_0^1 |g(\alpha; Q)|^2 \,\mathrm{d}\alpha \ll Q^{2+\varepsilon}. \tag{5.16}$$

Thus, on noting that our hypotheses on p ensure that $p^3P^2Q^2 \ll Q^5$, we deduce from (5.8), (5.9) and (5.13) that

$$\int_{\mathfrak{m}} |f(\alpha; P)g(\alpha p^3; Q)|^2 \,\mathrm{d}\alpha \ll Q^{5+\varepsilon} + p^3(I_1 + I_2), \tag{5.17}$$

where

$$I_1 = \int_{\mathfrak{n}} |F(\alpha)g(\alpha; Q)^2| \,\mathrm{d}\alpha \quad \text{and} \quad I_2 = \int_0^1 |G(\alpha)g(\alpha; Q)^2| \,\mathrm{d}\alpha. \tag{5.18}$$

We next exploit the congruence conditions implicit in the exponential sums $F(\alpha)$ and $G(\alpha)$. In (5.14) we substitute $x = u + hp^3$ and $y = v + kp^3$ to obtain

$$F(\alpha) = \sum_{\substack{0\leqslant|h|,|k|\leqslant H\\ (h,k)\neq(0,0)}} \sum_{\substack{(u,v)\in P\mathcal{B}\\ (u+hp^3, v+kp^3)\in P\mathcal{B}}} e(\alpha\Psi_{p^3}(u, v; h, k)),$$

where $H = \Omega P p^{-3}$, and Ψ_{p^3} is the polynomial defined in (2.5). It follows that $F(\alpha)$ takes a form to which lemma 2.2 is applicable, with $m = p^3$, and thus we deduce that

$$F(\alpha) \ll P^{3+\varepsilon}p^{-6} + P^{4+\varepsilon}p^{-6}K(\alpha), \tag{5.19}$$

where

$$K(\alpha) = \begin{cases} (q + Q^3\|q\alpha\|)^{-1}, & \text{when } q \leqslant P \text{ and } \|q\alpha\| \leqslant PQ^{-3},\\ 0, & \text{otherwise.} \end{cases} \tag{5.20}$$

Similarly, in (5.15) we substitute $x = u + hp^2$ and $y = v + kp^2$ to obtain

$$G(\alpha) = \sum_{\substack{0\leqslant|h|,|k|\leqslant H\\ (h,k)\neq(0,0)}} \sum_{\substack{(u,v)\in Q\mathcal{B}\\ (u+hp^2, v+kp^2)\in Q\mathcal{B}}} e(\alpha p^2\Psi_{p^2}(u, v; h, k)).$$

Thus $G(\alpha)$ also takes a form to which lemma 2.2 is applicable, but now with $m = p^2$, and, more importantly, with α replaced by αp^2. Hence, on writing

$$J(\alpha) = \begin{cases} (q + Q^3 p^{-2} \|q\alpha\|)^{-1}, & \text{when } q \leqslant Q \text{ and } \|q\alpha\| \leqslant p^2 Q^{-2}, \\ 0, & \text{otherwise,} \end{cases} \tag{5.21}$$

we deduce from lemma 2.2 that

$$G(\alpha) \ll P^{3+\varepsilon} p^{-7} + P^{4+\varepsilon} p^{-8} J(\alpha p^2). \tag{5.22}$$

On recalling (5.16)–(5.19) and (5.22), therefore, we may summarize our deliberations thus far by recording the estimate

$$\int_{\mathfrak{m}} |f(\alpha; P) g(\alpha p^3; Q)|^2 \, \mathrm{d}\alpha \ll Q^{5+\varepsilon} + P^{4+\varepsilon} p^{-3} I_3 + P^{4+\varepsilon} p^{-5} I_4, \tag{5.23}$$

where

$$I_3 = \int_{\mathfrak{n}} K(\alpha) |g(\alpha; Q)|^2 \, \mathrm{d}\alpha \quad \text{and} \quad I_4 = \int_0^1 J(\alpha p^2) |g(\alpha; Q)|^2 \, \mathrm{d}\alpha. \tag{5.24}$$

In order to complete the proof of the lemma, we have only to estimate I_3 and I_4. The estimation of I_4 is accomplished easily through the use of Schwarz's inequality, giving

$$I_4 \leqslant \left(\int_0^1 |g(\alpha; Q)|^4 \, \mathrm{d}\alpha \right)^{1/2} \left(\int_0^1 |J(\alpha p^2)|^2 \, \mathrm{d}\alpha \right)^{1/2}. \tag{5.25}$$

But by considering the underlying diophantine equations, it follows from lemma 4.1 that

$$\int_0^1 |g(\alpha; Q)|^4 \, \mathrm{d}\alpha \ll Q^{5+\varepsilon}. \tag{5.26}$$

Meanwhile, since $J(\alpha)$ is a periodic function of α with period 1, we deduce from (5.21) that

$$\begin{aligned} \int_0^1 |J(\alpha p^2)|^2 \, \mathrm{d}\alpha = \int_0^1 J(\alpha)^2 \, \mathrm{d}\alpha &\leqslant \sum_{1 \leqslant q \leqslant Q} q^{-1} \int_{-1/2}^{1/2} (1 + Q^3 p^{-2} |\beta|)^{-2} \, \mathrm{d}\beta \\ &\ll p^2 Q^{\varepsilon - 3}. \end{aligned} \tag{5.27}$$

On combining (5.25)–(5.27), therefore, we obtain

$$I_4 \ll Q^{1+\varepsilon} p. \tag{5.28}$$

It remains only to bound I_3. Recall the set of major arcs, $\mathfrak{N}(X)$, defined in § 3, and note that

$$I_3 = \int_{\mathfrak{n}} K(\alpha) |g(\alpha; Q)|^2 \, \mathrm{d}\alpha = \int_{\mathfrak{N}(P) \setminus \mathfrak{N}(R_0)} K(\alpha) |g(\alpha; Q)|^2 \, \mathrm{d}\alpha.$$

In view of (5.20), whenever $\alpha \in \mathfrak{N}(2X) \setminus \mathfrak{N}(X)$ one has $K(\alpha) \ll X^{-1}$. Then by dividing into dyadic intervals, one deduces that

$$I_3 \ll (\log P) \sup_{R_0 \leqslant X \leqslant P/2} X^{-1} \int_{\mathfrak{N}(2X) \setminus \mathfrak{N}(X)} |g(\alpha; Q)|^2 \, \mathrm{d}\alpha. \tag{5.29}$$

Next define the function $g^*(\alpha)$ by

$$g^*(\alpha) = \begin{cases} q^{-2}S(q,r)w(\alpha - r/q), & \text{when } \alpha \in \mathfrak{N}(q,r) \subseteq \mathfrak{N}(P), \\ 0, & \text{otherwise.} \end{cases} \tag{5.30}$$

Then in the notation introduced in (3.6), for each $\alpha \in \mathfrak{N}(P)$ one has

$$|g(\alpha;Q)|^2 \ll |g^*(\alpha)|^2 + |E(\alpha)|^2.$$

By lemma 3.1, therefore, we may conclude that whenever $1 \leqslant X \leqslant \frac{1}{2}Q^{3/2}$,

$$\int_{\mathfrak{N}(X)} |g(\alpha;Q)|^2 \,\mathrm{d}\alpha \ll \int_{\mathfrak{N}(X)} |g^*(\alpha)|^2 \,\mathrm{d}\alpha + X^{9/2}Q^{\varepsilon-3} + X^{1/2}Q^{1+\varepsilon}. \tag{5.31}$$

Moreover, when, for some positive number δ, one has $X \leqslant Q^{1-\delta}$, it follows from lemma 3.4 that one has the more precise estimate

$$\int_{\mathfrak{N}(X)} |g(\alpha;Q)|^2 \,\mathrm{d}\alpha \ll 1 + \int_{\mathfrak{N}(X)} |g^*(\alpha)|^2 \,\mathrm{d}\alpha. \tag{5.32}$$

Further, by (5.30), (3.5), and lemmata 3.3 and 2.6, in the notation used therein, whenever $1 \leqslant X \leqslant \frac{1}{2}Q^{3/2}$ one has

$$\begin{aligned} \int_{\mathfrak{N}(X)} |g^*(\alpha)|^2 \,\mathrm{d}\alpha &\ll Q^4 \sum_{1\leqslant q\leqslant X} \sum_{\substack{r=1 \\ (r,q)=1}}^{q} q^{-4}|S(q,r)|^2 \int_{-1/2}^{1/2} (1 + Q^3(\log Q)^{-2}|\beta|)^{-2} \,\mathrm{d}\beta \\ &\ll Q^{1+\varepsilon} \sum_{1\leqslant q\leqslant X} q^{1+\varepsilon} q_0^{-1} q_1^{-2} q_2^{-4/3} \ll Q^{1+\varepsilon}X^{\varepsilon}. \end{aligned} \tag{5.33}$$

On combining (5.29) and (5.31)–(5.33), therefore, we may finally conclude that

$$\begin{aligned} I_3 &\ll P^{\varepsilon} \sup_{R_0 \leqslant X \leqslant Q^{1-\varepsilon}} (X^{-1} + QX^{-1}) + P^{\varepsilon} \sup_{Q^{1-\varepsilon} < X \leqslant P} (X^{7/2}Q^{-3} + X^{-1/2}Q) \\ &\ll P^{\varepsilon}(p^3QR^{-1} + P^{7/2}Q^{-3} + Q^{1/2}). \end{aligned}$$

Since, by hypothesis, we have $p \leqslant P^{1/10}$ and $R \geqslant (\frac{1}{2}p)^4$, we arrive at the estimate

$$I_3 \ll Q^{1+\varepsilon}p^{-1}. \tag{5.34}$$

The proof of the lemma is completed on combining (5.23), (5.28) and (5.34). ■

6. The proof of theorem 1.1: the lower bound

Our preliminary campaigning on the minor arcs now complete, we launch our assault on the proof of the lower bound recorded in theorem 1.1. We first dispose of a simple exceptional case. Suppose that $\Phi(x,y)$ is a binary cubic form with integer coefficients having non-zero discriminant D, and suppose that $-3D$ is a square. Then it is well known (see, for example, lemma 18 of Chowla & Davenport (1961)) that there exist rational numbers A and B, and linearly independent linear forms $X = X(x,y)$ and $Y = Y(x,y)$ with rational coefficients, such that $\Phi(x,y) = AX^3 + BY^3$. Suppose next that for each j with $1 \leqslant j \leqslant 4$, the binary cubic form Φ_j, with discriminant D_j, satisfies the condition that $-3D_j$ is a square. Then by the above discussion, there exists a positive number δ and integers c_i $(1 \leqslant i \leqslant 8)$ depending at most on $\boldsymbol{\Phi}$, such

that the set of solutions counted by $\mathcal{N}(P;\boldsymbol{\Phi})$ is in bijective correspondence with a subset of the solutions of the diophantine equation

$$c_1x_1^3 + c_2x_2^3 + \cdots + c_8x_8^3 = 0, \tag{6.1}$$

with $|x_i| \leqslant \delta P$ $(1 \leqslant i \leqslant 8)$. Moreover, this correspondence is defined by a non-degenerate linear transformation with rational coefficients, so that a little thought reveals that a lower bound for $\mathcal{N}(P;\boldsymbol{\Phi})$ is provided by the number of solutions of (6.1) subject to congruence conditions on the x_i, and with the x_i confined to a convex subset of $[-\delta P, \delta P]^8$ with volume $\gg P^8$. We note that a further consequence of the non-degeneracy of the aforementioned transformation is the non-trivial local solubility of equation (1.1) in this case. For the non-trivial p-adic solubility of equation (6.1), for each prime p, is immediate from Lewis (1957*b*), and any non-trivial p-adic solution of the latter equation yields a non-trivial p-adic solution of (1.1). Thus the work of Vaughan (1986), combined with standard major arc techniques from the Hardy–Littlewood method (see, for example, Vaughan 1997), can be applied to yield an asymptotic formula for $\mathcal{N}(P;\boldsymbol{\Phi})$ of the shape

$$\mathcal{N}(P;\boldsymbol{\Phi}) \sim \kappa P^5,$$

for a suitable positive number κ, and indeed the lower bound $\mathcal{N}(P;\boldsymbol{\Phi}) \gg P^5$ is simpler still to obtain. Consequently, in the above special case, the lower bound of theorem 1.1 follows easily without recourse to the main techniques of this paper.

In view of the above deliberations, we may suppose that there is an i for which $-3D_i$ is not a square. Henceforth we suppose that $-3D_2$ is not a perfect square. Let δ be a fixed number with $0 < \delta < \frac{1}{10}$, and choose an admissible point (ξ_1, η_1) for Φ_1. Plainly, we may also choose real numbers ξ_j, η_j $(2 \leqslant j \leqslant 4)$ satisfying

$$\Phi_1(\xi_1, \eta_1) + \cdots + \Phi_4(\xi_4, \eta_4) = 0. \tag{6.2}$$

Moreover, the non-vanishing of the discriminants D_j ensures that we may make a non-singular such choice for $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$, whence we may suppose that for some i with $1 \leqslant i \leqslant 4$ one has

$$\frac{\partial \Phi_i}{\partial \xi}(\xi_i, \eta_i) \neq 0 \quad \text{or} \quad \frac{\partial \Phi_i}{\partial \eta}(\xi_i, \eta_i) \neq 0. \tag{6.3}$$

Let τ be a sufficiently small positive number, and define the boxes

$$\mathcal{B}_j = \{(\xi, \eta) : |\xi - \xi_j| \leqslant \tau \text{ and } |\eta - \eta_j| \leqslant \tau\} \quad (1 \leqslant j \leqslant 4). \tag{6.4}$$

Choose a prime p with $P^\delta < p \leqslant 2P^\delta$, write $Q = P/p$, and consider the exponential sums

$$g(\alpha) = \sum_{(x,y)\in\mathbb{Z}^2} \Gamma_Q(x, y) e(\alpha\Phi_1(x, y)), \tag{6.5}$$

and

$$f_j(\alpha) = \sum_{(x,y)\in P\mathcal{B}_j} e(\alpha\Phi_j(x, y)) \quad (2 \leqslant j \leqslant 4).$$

By orthogonality, the integral

$$\mathcal{N}_1 = \int_0^1 g(\alpha p^3) f_2(\alpha) f_3(\alpha) f_4(\alpha)\, \mathrm{d}\alpha \tag{6.6}$$

counts the number of solutions of the diophantine equation

$$p^3\Phi_1(x_1,y_1)+\Phi_2(x_2,y_2)+\Phi_3(x_3,y_3)+\Phi_4(x_4,y_4)=0, \tag{6.7}$$

with $(x_1,y_1)\in\mathbb{Z}^2$ and $(x_j,y_j)\in P\mathcal{B}_j$ $(2\leqslant j\leqslant 4)$, and with each solution $(\boldsymbol{x},\boldsymbol{y})$ counted with weight $\Gamma_Q(x_1,y_1)$. Hence, in view of (3.1) and (3.2), there is a positive number Ω, depending at most on our choices for $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$, such that

$$\mathcal{N}(\Omega P;\boldsymbol{\Phi})\geqslant\mathcal{N}_1. \tag{6.8}$$

In the remainder of this section we will establish the lower bound

$$\mathcal{N}_1\gg P^{5-2\delta}, \tag{6.9}$$

so that since $\delta>0$ can be taken as small as we please, the lower bound in theorem 1.1 will follow immediately from (6.8) and (6.9).

We establish (6.9) by using the Hardy–Littlewood method. Let $\mathfrak{M}$ denote the union of the intervals

$$\mathfrak{M}(q,r)=\{\alpha\in[0,1):|q\alpha-r|\leqslant P^{4\delta-3}\}$$

with $0\leqslant r\leqslant q\leqslant P^{4\delta}$ and $(r,q)=1$, and let $\mathfrak{m}=[0,1)\setminus\mathfrak{M}$. By Schwarz's inequality one has

$$\int_{\mathfrak{m}}|g(\alpha p^3)f_2(\alpha)f_3(\alpha)f_4(\alpha)|\,\mathrm{d}\alpha\leqslant J^{1/2}J_3^{1/4}J_4^{1/4}, \tag{6.10}$$

where

$$J=\int_{\mathfrak{m}}|f_2(\alpha)g(p^3\alpha)|^2\,\mathrm{d}\alpha\quad\text{and}\quad J_i=\int_0^1|f_i(\alpha)|^4\,\mathrm{d}\alpha\quad(i=3,4).$$

But by lemmata 5.2 and 4.1, respectively, one has $J\ll Q^{5+\varepsilon}$, and $J_i\ll P^{5+\varepsilon}$ $(i=3,4)$. Then the definition of $\mathfrak{m}$, together with (6.6) and (6.10), implies that

$$\mathcal{N}_1=\int_{\mathfrak{M}}g(\alpha p^3)f_2(\alpha)f_3(\alpha)f_4(\alpha)\,\mathrm{d}\alpha+O(P^{5-(5/2)\delta+\varepsilon}). \tag{6.11}$$

The evaluation of the major arc integral is routine, and thus we will be brief. Let

$$S_j(q,r)=\sum_{x=1}^{q}\sum_{y=1}^{q}e\left(\frac{r}{q}\Phi_j(x,y)\right)\quad(1\leqslant j\leqslant 4), \tag{6.12}$$

$$v_j(\beta;P)=\iint_{P\mathcal{B}_j}e(\beta\Phi_j(\xi,\eta))\,\mathrm{d}\xi\mathrm{d}\eta\quad(2\leqslant j\leqslant 4),$$

and

$$w(\beta;Q)=\iint_{\mathbb{R}^2}\Gamma_Q(\xi,\eta)e(\beta\Phi_1(\xi,\eta))\,\mathrm{d}\xi\mathrm{d}\eta. \tag{6.13}$$

When $r\in\mathbb{Z}$ and $q\in\mathbb{N}$ satisfy $0\leqslant r\leqslant q\leqslant P^{4\delta}$ and $(r,q)=1$, define $f_j^*(\alpha)$ and $g_p^*(\alpha)$ for $|q\alpha-r|\leqslant P^{4\delta-3}$ by

$$f_j^*(\alpha)=q^{-2}S_j(q,r)v_j(\alpha-r/q;P)\quad(2\leqslant j\leqslant 4), \tag{6.14}$$

and

$$g_p^*(\alpha)=q^{-2}S_1(q,rp^3)w(p^3(\alpha-r/q);Q). \tag{6.15}$$

Then by lemmata 2.3 and 3.4 together with (3.6), whenever $\alpha \in \mathfrak{M}(q,r) \subseteq \mathfrak{M}$, one has

$$g(\alpha p^3) - g_p^*(\alpha) \ll 1 \quad \text{and} \quad f_j(\alpha) - f_j^*(\alpha) \ll P^{1+\varepsilon} q^{1/2} \quad (2 \leqslant j \leqslant 4).$$

In view of lemma 2.6, however, when $\alpha \in \mathfrak{M}(q,r) \subseteq \mathfrak{M}$, one has $f_j^*(\alpha) \ll P^2 q^{\varepsilon - 1/2}$ $(2 \leqslant j \leqslant 4)$, whence when $2 \leqslant i, j \leqslant 4$ and $i \neq j$,

$$\sup_{\alpha \in \mathfrak{M}} |f_i^*(\alpha)(f_j(\alpha) - f_j^*(\alpha))| \ll P^{3+\varepsilon}.$$

Thus it is readily confirmed that

$$g(\alpha p^3) f_2(\alpha) f_3(\alpha) f_4(\alpha) - g_p^*(\alpha) f_2^*(\alpha) f_3^*(\alpha) f_4^*(\alpha) \ll P^{7+\varepsilon}.$$

Since the measure of $\mathfrak{M}$ is $O(P^{8\delta-3})$, we deduce that

$$\int_{\mathfrak{M}} g(\alpha p^3) f_2(\alpha) f_3(\alpha) f_4(\alpha)\, \mathrm{d}\alpha - \int_{\mathfrak{M}} g_p^*(\alpha) f_2^*(\alpha) f_3^*(\alpha) f_4^*(\alpha)\, \mathrm{d}\alpha \ll P^{4+8\delta+\varepsilon}. \quad (6.16)$$

By a change of variable, one has $w(p^3\beta; Q) = p^{-2} w(\beta; P)$. Thus, on substituting from (6.14) and (6.15), we deduce that

$$\int_{\mathfrak{M}} g_p^*(\alpha) f_2^*(\alpha) f_3^*(\alpha) f_4^*(\alpha)\, \mathrm{d}\alpha = p^{-2} \sum_{1 \leqslant q \leqslant P^{4\delta}} S_p(q) J(q), \quad (6.17)$$

where

$$S_p(q) = q^{-8} \sum_{\substack{r=1 \\ (r,q)=1}}^{q} S_1(q, rp^3) S_2(q,r) S_3(q,r) S_4(q,r) \quad (6.18)$$

and

$$J(q) = \int_{-q^{-1}P^{4\delta-3}}^{q^{-1}P^{4\delta-3}} w(\beta; P) v_2(\beta; P) v_3(\beta; P) v_4(\beta; P)\, \mathrm{d}\beta.$$

We are able to dispose of the analysis of the singular integral $J(q)$ swiftly by making another change of variables. Thus a straightforward application of lemma 2.7 yields

$$\begin{aligned} J(q) &= P^5 \int_{-q^{-1}P^{4\delta}}^{q^{-1}P^{4\delta}} w(\beta; 1) v_2(\beta; 1) v_3(\beta; 1) v_4(\beta; 1)\, \mathrm{d}\beta \\ &= P^5 (J + O(qP^{-4\delta})), \end{aligned} \quad (6.19)$$

where

$$J = \int_{-\infty}^{\infty} w(\beta; 1) v_2(\beta; 1) v_3(\beta; 1) v_4(\beta; 1)\, \mathrm{d}\beta.$$

Moreover, when τ is sufficiently small, bearing in mind (6.2)–(6.4), a standard application of Fourier's integral formula (see, for example, lemma 6.2 of Davenport (1959)) shows that $J > 0$. Thus (6.17) and (6.19) imply that

$$\int_{\mathfrak{M}} g_p^*(\alpha) f_2^*(\alpha) f_3^*(\alpha) f_4^*(\alpha)\, \mathrm{d}\alpha = P^5 p^{-2} (J \mathfrak{S}_p(P^{4\delta}) + O(\mathfrak{T}(P^{4\delta}))), \quad (6.20)$$

where

$$\mathfrak{S}(X) = \sum_{1 \leqslant q \leqslant X} S_p(q) \quad \text{and} \quad \mathfrak{T}(X) = X^{-1} \sum_{1 \leqslant q \leqslant X} q |S_p(q)|. \quad (6.21)$$

We now turn our attention to the task of estimating $S_p(q)$. For each $q \in \mathbb{N}$ there is a unique decomposition $q = q_0q_1q_2$, where q_0, q_1 and q_2 are pairwise coprime, and where q_1 is cube-free, q_2 is cube-full, $(q_1q_2, 6D_2D_3D_4) = 1$, and whenever $p|q_0$ one has $p|6D_2D_3D_4$. By estimating $S_1(q, rp^3)$ trivially, and applying lemma 2.6 for the remaining terms, we obtain from (6.18) the estimate

$$S_p(q) \ll q^{1+\varepsilon}q_0^{-3/2}q_1^{-3}q_2^{-2}.$$

Consequently, bearing in mind our notational devices, one has

$$\begin{aligned} X^{1/2-2\varepsilon}\sum_{q>X}|S_p(q)| &\ll \sum_{q_0q_1q_2>X} q_0^{-\varepsilon}q_1^{-3/2}q_2^{-1/2} \\ &\ll \prod_{p|6D_2D_3D_4}(1-p^{-\varepsilon})^{-1}\sum_{\substack{q_1 \text{ cube-free}\\ q_2 \text{ cube-full}}} q_1^{-3/2}q_2^{-1/2} \\ &\ll 1, \end{aligned}$$

whence

$$\sum_{q>X}|S_p(q)| \ll X^{\varepsilon-1/2}. \tag{6.22}$$

Similarly,

$$X^{-1/2-2\varepsilon}\sum_{1\leqslant q\leqslant X} q|S_p(q)| \ll \sum_{q_0q_1q_2\leqslant X} q_0^{-\varepsilon}q_1^{-3/2}q_2^{-1/2} \ll 1,$$

whence

$$\sum_{1\leqslant q\leqslant X} q|S_q(p)| \ll X^{1/2+\varepsilon}. \tag{6.23}$$

It follows in particular that the singular series

$$\mathfrak{S}_p = \sum_{q=1}^{\infty} S_p(q) \tag{6.24}$$

converges absolutely. Moreover, on applying the estimates (6.22) and (6.23) to (6.21), we obtain

$$\mathfrak{S}(X) = \mathfrak{S}_p + O(X^{\varepsilon-1/2}) \quad \text{and} \quad \mathfrak{T}(X) \ll X^{\varepsilon-1/2}, \tag{6.25}$$

whence by (6.20) we deduce that

$$\int_{\mathfrak{M}} g_p^*(\alpha)f_2^*(\alpha)f_3^*(\alpha)f_4^*(\alpha)\,\mathrm{d}\alpha = P^5p^{-2}(J\mathfrak{S}_p + O(P^{\varepsilon-2\delta})). \tag{6.26}$$

On combining (6.8), (6.11), (6.16) and (6.26), we may conclude thus far that

$$\mathcal{N}(\Omega P; \boldsymbol{\Phi}) \gg J\mathfrak{S}_pP^{5-2\delta},$$

and thus, on taking δ to be a sufficiently small positive number, the lower bound of theorem 1.1 will follow from the positivity of J, provided that we confirm that $\mathfrak{S}_p \gg 1$.

We now analyse the singular series at a modest level. We begin by noting that (2.14) and (6.18) together imply that $S_p(q)$ is a multiplicative function of q. Thus, in view of the absolute convergence of the series recorded in (6.24), one has

$$\mathfrak{S}_p = \prod_{\varpi} \sum_{h=0}^{\infty} S_p(\varpi^h),$$

where the product is over prime numbers ϖ. Then by (6.22), on taking ϖ_0 to be a sufficiently large constant depending at most on the coefficients of the Φ_i, one has

$$\mathfrak{S}_p = \Big(\prod_{\varpi \leqslant \varpi_0} \sum_{h=0}^{\infty} S_p(\varpi^h) \Big) (1 + O(\varpi_0^{-1/4})) \geqslant \frac{1}{2} \prod_{\varpi \leqslant \varpi_0} \sum_{h=0}^{\infty} S_p(\varpi^h). \tag{6.27}$$

When P is sufficiently large one has $p > \varpi_0$, and moreover when $\varpi \neq p$ a simple substitution reveals that $S_1(\varpi^h, rp^3) = S_1(\varpi^h, r)$. Thus we deduce from (6.18) and (6.27) that

$$\mathfrak{S}_p \geqslant \frac{1}{2} \prod_{\varpi \leqslant \varpi_0} \chi(\varpi), \tag{6.28}$$

where

$$\chi(\varpi) = \sum_{h=0}^{\infty} \varpi^{-8h} \sum_{\substack{r=1 \\ (r,\varpi)=1}}^{\varpi^h} S_1(\varpi^h, r) S_2(\varpi^h, r) S_3(\varpi^h, r) S_4(\varpi^h, r).$$

But standard methods (see, for example, the treatment provided by Davenport (1959)) show that, for each prime number ϖ, the existence of a non-singular ϖ-adic solution of equation (1.1) implies that $\chi(\varpi) > 0$. Then in view of (6.28), we obtain the desired conclusion that $\mathfrak{S}_p$ is positive, and bounded uniformly away from zero, provided only that equation (1.1) possesses a non-singular solution in every ϖ-adic field.

To complete our proof of theorem 1.1, we now briefly sketch how to prove that (1.1) indeed possesses a non-singular solution in every ϖ-adic field, using the argument suggested by that completing Chowla & Davenport (1961). We note first that since the two first partial derivatives of a binary cubic form with non-zero discriminant cannot vanish unless both variables vanish, it suffices to prove the existence of a non-trivial solution of (1.1) in every ϖ-adic field. One observes next that an integral binary cubic form with non-zero discriminant is equivalent, under a linear transformation over a quadratic field extension of $\mathbb{Q}_\varpi$, to a diagonal form with coefficients in the latter field extension. Consequently, equation (1.1) is equivalent to a diagonal equation defined over a field K_ϖ which arises from a succession of quadratic extensions of $\mathbb{Q}_\varpi$. Given the existence of a non-trivial solution to the latter equation in K_ϖ, which is guaranteed by the principal conclusion of Lewis (1957*b*), one may employ an argument of Lewis (1957*a*) to pull this solution back, through the tower of quadratic extensions, to a non-trivial ϖ-adic solution of (1.1). This completes our sketch of the local solubility behaviour relevant to our argument, and hence also the proof of theorem 1.1.

7. The proof of theorem 1.2

Having completed our analysis of sums of four binary cubic forms, we now investigate the representation of integers as the sum of two such forms. For $j = 1, 2$, let $\Phi_j(x, y)$ denote an integral binary cubic form with non-zero discriminant D_j, and consider equation (1.3). As in the previous section, we dispose first of the cases in which $-3D_j$ is a perfect square for $j = 1$ and 2. In this situation the forms Φ_j $(j = 1, 2)$ each diagonalize via a non-degenerate rational transformation. It follows that there exist positive integers B and J, and integers b_{ij} and c_i $(1 \leqslant i \leqslant 4, 1 \leqslant j \leqslant J)$, all depending at most on Φ_1 and Φ_2, such that an integer n is represented in the form (1.3) if and only if n is represented in the form

$$c_1 y_1^3 + c_2 y_2^3 + c_3 y_3^3 + c_4 y_4^3 = n, \tag{7.1}$$

with the variables satisfying the congruence conditions

$$y_i \equiv b_{ij} \pmod{B} \quad (1 \leqslant i \leqslant 4) \tag{7.2}$$

for some j with $1 \leqslant j \leqslant J$. The latter problem is within the compass of standard circle method machinery. Indeed the methods of Davenport (1939) suffice to show that each integer n with $|n| \leqslant X$, for which the local solubility conditions are satisfied, has a representation of the form (7.1) subject to (7.2), with at most $O(X^{29/30+\varepsilon})$ exceptions. We may omit the (standard) details in the interests of concision.

Consider now a large positive number N, and a sufficiently small positive number τ. We consider an integer n satisfying $(1 - \tau)N < n \leqslant N$. In view of the above discussion we may suppose that for $i = 1$ or 2 the discriminant $-3D_i$ is not a square. Henceforth we suppose that $-3D_2$ is not a square. Choose an admissible point (ξ_1, η_1) for Φ_1. Plainly, we may also choose real numbers ξ_2, η_2 satisfying

$$\Phi_1(\xi_1, \eta_1) + \Phi_2(\xi_2, \eta_2) = 1. \tag{7.3}$$

Moreover, the non-vanishing of the discriminants D_j ensures that we may make a non-singular choice for $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$, whence we may suppose that for at least one of $i = 1$ or $i = 2$ one has

$$\frac{\partial \Phi_i}{\partial \xi}(\xi_i, \eta_i) \neq 0 \quad \text{or} \quad \frac{\partial \Phi_i}{\partial \eta}(\xi_i, \eta_i) \neq 0. \tag{7.4}$$

Define the box

$$\mathcal{B} = \{(\xi, \eta) : |\xi - \xi_2| \leqslant \tau \text{ and } |\eta - \eta_2| \leqslant \tau\}. \tag{7.5}$$

Let $P = N^{1/3}$, and consider the exponential sum

$$f(\alpha) = \sum_{(x,y) \in P\mathcal{B}} e(\alpha \Phi_2(x, y)).$$

Choose a prime p with $P^{1/10} < p \leqslant 2P^{1/10}$, write $Q = P/p$, and define $g(\alpha)$ as in (6.5). Further, define

$$\rho(n) = \int_0^1 f(\alpha) g(p^3 \alpha) e(-\alpha n) \, \mathrm{d}\alpha. \tag{7.6}$$

Then by orthogonality one finds that $\rho(n)$ counts solutions of equation (1.3) with a certain non-negative weight. We aim to show that for each large N, one has $\rho(n) \geqslant 1$

for all but $O(N^{209/210+\varepsilon})$ of the integers $n \in \mathcal{W}$ with $(1-\tau)N < n \leqslant N$, so long as τ is sufficiently small. Theorem 1.2 will follow from the latter conclusion so long as we are able to show that $\mathcal{W}$ has positive density.

Before describing the first step in our analysis, we require some notation. We define $S_j(q,r)$ $(j=1,2)$ as in (6.12). We also define $w(\beta;Q)$ as in (6.13), and write

$$v(\beta;P) = \int\int_{P\mathcal{B}} e(\beta\Phi_2(\xi,\eta))\,\mathrm{d}\xi\mathrm{d}\eta. \tag{7.7}$$

We then define the singular integral $J(n)$ by

$$J(n) = \int_{-\infty}^{\infty} v(\beta;P)w(\beta;P)e(-\beta n)\,\mathrm{d}\beta. \tag{7.8}$$

Finally, we define the singular series $\mathfrak{S}(n)$ by

$$\mathfrak{S}(n) = \sum_{q=1}^{\infty} q^{-4}\mathfrak{A}(q,n), \tag{7.9}$$

where

$$\mathfrak{A}(q,n) = \sum_{\substack{r=1\\(r,q)=1}}^{q} S_1(q,r)S_2(q,r)e(-rn/q). \tag{7.10}$$

Lemma 7.1. *In the notation introduced above, one has*

$$\sum_{1\leqslant n\leqslant N} |\rho(n) - p^{-2}J(n)\mathfrak{S}(n)|^2 \ll Q^{5+\varepsilon}.$$

Proof. Before launching into the proof proper, we arm ourselves with a little notation. For the sake of concision, write $L = P^{1/10}$. When $r \in \mathbb{Z}$ and $q \in \mathbb{N}$, define the major arc $\mathfrak{M}(q,r)$ by

$$\mathfrak{M}(q,r) = \{\alpha \in [0,1) : |q\alpha - r| \leqslant LP^{-3}\}.$$

We take $\mathfrak{M}$ to be the union of the intervals $\mathfrak{M}(q,r)$ with $0 \leqslant r \leqslant q \leqslant L$ and $(r,q) = 1$. Let $\mathfrak{m} = [0,1) \setminus \mathfrak{M}$. Next, when $r \in \mathbb{Z}$ and $q \in \mathbb{N}$, define the wider major arc $\mathfrak{N}(q,r)$ by

$$\mathfrak{N}(q,r) = \{\alpha \in [0,1) : |q\alpha - r| \leqslant L^4P^{-3}\}.$$

We take $\mathfrak{N}$ to be the union of the intervals $\mathfrak{N}(q,r)$ with $0 \leqslant r \leqslant q \leqslant L^4$ and $(r,q) = 1$. Let $\mathfrak{n} = [0,1) \setminus \mathfrak{N}$. Finally, when $\mathfrak{B} \subseteq [0,1)$, put

$$\rho(n,\mathfrak{B}) = \int_{\mathfrak{B}} f(\alpha)g(p^3\alpha)e(-\alpha n)\,\mathrm{d}\alpha.$$

We first estimate $\rho(n,\mathfrak{m})$ in mean square. By Bessel's inequality we have

$$\sum_{n\in\mathbb{Z}} |\rho(n,\mathfrak{m})|^2 \leqslant \int_{\mathfrak{m}} |f(\alpha)g(p^3\alpha)|^2\,\mathrm{d}\alpha. \tag{7.11}$$

In order to estimate the latter mean value, we start by dissecting $\mathfrak{m}$ as $\mathfrak{n} \cup (\mathfrak{N} \setminus \mathfrak{M})$, and note that lemma 5.2 yields the estimate

$$\int_{\mathfrak{n}} |f(\alpha)g(p^3\alpha)|^2\,\mathrm{d}\alpha \ll Q^{5+\varepsilon}. \tag{7.12}$$

We next conduct a pruning operation, establishing a similar estimate to that provided by (7.12) with $\mathfrak{N} \setminus \mathfrak{M}$ in place of $\mathfrak{n}$. When $r \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $0 \leqslant r \leqslant q \leqslant L^4$ and $(r,q) = 1$, define $f^*(\alpha)$ for $|q\alpha - r| \leqslant L^4P^{-3}$ by

$$f^*(\alpha) = q^{-2}S_2(q,r)v\left(\alpha - \frac{r}{q}; P\right). \tag{7.13}$$

Also, define $g_p^*(\alpha)$ as in (6.15). Then, as in the argument of §6 leading to (6.16), when $\alpha \in \mathfrak{N}(q,r) \subseteq \mathfrak{N}$ one has

$$g(p^3\alpha) - g_p^*(\alpha) \ll 1 \quad \text{and} \quad f(\alpha) - f^*(\alpha) \ll P^{1+\varepsilon}q^{1/2}.$$

Hence, when $\alpha \in \mathfrak{N}$,

$$|f(\alpha)|^2 \ll |f^*(\alpha)|^2 + P^{2+\varepsilon}L^4 \ll |f^*(\alpha)|^2 + P^{12/5+\varepsilon},$$

and similarly,

$$|g(p^3\alpha)|^2 \ll |g_p^*(\alpha)|^2 + 1.$$

Consequently, trivial bounds for $f^*(\alpha)$ and $g_p^*(\alpha)$ now suffice to confirm that whenever $\alpha \in \mathfrak{N}$, one has

$$|f(\alpha)g(p^3\alpha)|^2 \ll |f^*(\alpha)g_p^*(\alpha)|^2 + P^{12/5+\varepsilon}Q^4.$$

The measure of $\mathfrak{N}$ is $O(L^8P^{-3})$, and thus we deduce that

$$\int_{\mathfrak{N}\setminus\mathfrak{M}} |f(\alpha)g(p^3\alpha)|^2\,\mathrm{d}\alpha \ll \int_{\mathfrak{N}\setminus\mathfrak{M}} |f^*(\alpha)g_p^*(\alpha)|^2\,\mathrm{d}\alpha + P^{1/5+\varepsilon}Q^4. \tag{7.14}$$

We next recall that $w(p^3\beta; Q) = p^{-2}w(\beta; P)$, and estimate $w(\beta; P)$ by lemma 3.3. Then by (6.13), (6.15) and (7.13), we obtain

$$\int_{\mathfrak{N}\setminus\mathfrak{M}} |f^*(\alpha)g_p^*(\alpha)|^2\,\mathrm{d}\alpha \ll P^{4+\varepsilon}Q^4 \sum_{1\leqslant q\leqslant L^4} B_p(q) \int_{\mathfrak{P}(q)} (1 + P^3|\beta|)^{-2}\,\mathrm{d}\beta, \tag{7.15}$$

where

$$B_p(q) = \sum_{\substack{r=1\\(r,q)=1}}^{q} q^{-8}|S_1(q, rp^3)S_2(q,r)|^2,$$

and

$$\mathfrak{P}(q) = \begin{cases} \mathbb{R}, & \text{when } q > L, \\ \mathbb{R} \setminus [-q^{-1}LP^{-3}, q^{-1}LP^{-3}], & \text{when } 1 \leqslant q \leqslant L. \end{cases}$$

Note that $B_p(q)$ is a multiplicative function of q. For each $q \in \mathbb{N}$ there is a unique decomposition $q = q_0q_1q_2$, where q_0, q_1 and q_2 are pairwise coprime, and where q_1 is cube-free, q_2 is cube-full, $(q_1q_2, 6D_1D_2) = 1$, and whenever $p|q_0$ one has $p|6D_1D_2$. When $p \nmid q$, it is a consequence of lemma 2.6 that

$$B_p(q) \ll q^{1+\varepsilon}q_0^{-2}q_1^{-4}q_2^{-8/3}. \tag{7.16}$$

Meanwhile, when q is a power of p we estimate $S_1(q, rp^3)$ trivially. Thus, since $p > 6D_1D_2$ when P is large, it follows from lemmata 2.4 and 2.5 that

$$B_p(p) \ll p^{-1}, \quad B_p(p^2) \ll p^{-2} \quad \text{and} \quad B_p(p^l) \ll p^{-l/3} \quad (l \geqslant 3). \tag{7.17}$$

By considering the formal Euler product, for example, one readily deduces from (7.16) that the series

$$\sum_{\substack{q=1\\(q,p)=1}}^{\infty} q^{\theta} B_p(q)$$

is absolutely convergent for $\theta < 1$. Then on noting that $p > L$, and making use of (7.17), we deduce that

$$\sum_{1 \leqslant q \leqslant L} qB_p(q) \ll L^{\varepsilon} \quad \text{and} \quad \sum_{L < q \leqslant L^4} B_p(q) \ll L^{\varepsilon - 1}. \tag{7.18}$$

On substituting (7.18) into (7.15), therefore, we arrive at the estimate

$$\int_{\mathfrak{N} \setminus \mathfrak{M}} |f^*(\alpha) g_p^*(\alpha)|^2 \, \mathrm{d}\alpha \ll P^{1+\varepsilon} Q^4 \sum_{1 \leqslant q \leqslant L^4} B_p(q) \min\{1, q/L\} \ll Q^{5+\varepsilon}. \tag{7.19}$$

On combining (7.11), (7.12), (7.14) and (7.19), we may conclude thus far that

$$\sum_{n \in \mathbb{Z}} |\rho(n, \mathfrak{m})|^2 \ll Q^{5+\varepsilon}. \tag{7.20}$$

Our next step is to compare $\rho(n, \mathfrak{M})$ with

$$\rho^*(n) = \int_{\mathfrak{M}} f^*(\alpha) g_p^*(\alpha) e(-\alpha n) \, \mathrm{d}\alpha. \tag{7.21}$$

By lemmata 2.3 and 3.4, whenever $\alpha \in \mathfrak{M}$ one has

$$f(\alpha) - f^*(\alpha) \ll P^{1+\varepsilon} L^{1/2} \quad \text{and} \quad g(\alpha p^3) - g_p^*(\alpha) \ll 1.$$

Hence, for $\alpha \in \mathfrak{M}$,

$$f(\alpha) g(\alpha p^3) - f^*(\alpha) g_p^*(\alpha) \ll P^{1+\varepsilon} Q^2 L^{1/2}.$$

Since the measure of $\mathfrak{M}$ is $O(L^2 P^{-3})$, therefore, another application of Bessel's inequality yields

$$\begin{aligned} \sum_{n \in \mathbb{Z}} |\rho(n, \mathfrak{M}) - \rho^*(n)|^2 &\leqslant \int_{\mathfrak{M}} |f(\alpha) g(\alpha p^3) - f^*(\alpha) g_p^*(\alpha)|^2 \, \mathrm{d}\alpha \\ &\ll (P^{1+\varepsilon} Q^2 L^{1/2})^2 P^{-3} L^2 \ll Q^5, \end{aligned}$$

and thus it follows from (7.20) that

$$\sum_{n \in \mathbb{Z}} |\rho(n) - \rho^*(n)|^2 \ll Q^{5+\varepsilon}. \tag{7.22}$$

We now extract a main term from $\rho^*(n)$. In view of our choice for p, whenever $1 \leqslant q \leqslant L$ one has $p \nmid q$, and hence $S_1(q, rp^3) = S_1(q, r)$. It therefore follows from (6.15) and (7.13) that (7.21) may be rewritten in the shape

$$\rho^*(n) = p^{-2} \sum_{1 \leqslant q \leqslant L} q^{-4} \mathfrak{A}(q, n) \int_{-q^{-1} L P^{-3}}^{q^{-1} L P^{-3}} v(\beta; P) w(\beta; P) e(-\beta n) \, \mathrm{d}\beta, \tag{7.23}$$

where $\mathfrak{A}(q,n)$ is defined by (7.10). Further progress on the estimation of (7.23) now depends on an estimate for $\mathfrak{A}(q,n)$ superior to that which follows immediately from lemma 2.6. We presently establish the estimate

$$\mathfrak{A}(q,n) \ll q^{1/2+\varepsilon} q_0^3 q_1^2 q_2^{8/3} (q,n)^{1/2}, \tag{7.24}$$

where $q = q_0 q_1 q_2$ is the decomposition introduced in the preamble to (7.16). The proof of (7.24) depends on a suitable transformation of (7.10). Let t be an integer with $(t,q) = 1$. Then since $t^3 \Phi_j(x,y) = \Phi_j(tx,ty)$ $(j = 1,2)$, a simple substitution in (6.12) yields $S_j(q,r) = S_j(q,t^3 r)$. On substituting the latter into (7.10), and substituting also r for occurrences of rt^3, we deduce that $\mathfrak{A}(q,n) = \mathfrak{A}(q,l^3 n)$, where l satisfies $lt \equiv 1 \pmod q$. Consequently, on summing over the values of l with $(l,q) = 1$, we deduce that

$$\phi(q)\mathfrak{A}(q,n) = \sum_{\substack{l=1 \\ (l,q)=1}}^{q} \mathfrak{A}(q,l^3 n) = \sum_{\substack{r=1 \\ (r,q)=1}}^{q} S_1(q,r) S_2(q,r) U(q,-rn), \tag{7.25}$$

where

$$U(q,b) = \sum_{\substack{l=1 \\ (l,q)=1}}^{q} e(bl^3/q).$$

Plainly,

$$U(q,b) = (q,b) U\left(\frac{q}{(q,b)}, \frac{b}{(q,b)}\right).$$

Moreover, lemma 1.3 of Hua (1938) shows that whenever $(q,b) = 1$, one has $U(q,b) \ll q^{1/2+\varepsilon}$. We therefore deduce that whenever $(r,q) = 1$, one has the estimate

$$U(q,rn) \ll q^{1/2+\varepsilon} (q,n)^{1/2},$$

and thus (7.24) follows from (7.25) together with lemma 2.6.

For later use, we note also that $\mathfrak{A}(q,n)$ is a multiplicative function of q. Further, for primes ϖ with $\varpi \nmid b$, lemma 1.2 of Hua (1938) shows that $U(\varpi^h, b) = 0$ for $h > \gamma$, where $\gamma = 1$ when $\varpi \neq 3$, and $\gamma = 2$ when $\varpi = 3$. For a given non-zero integer n, let $\nu(\varpi)$ be the exact power of ϖ dividing n. Then it follows that when $(r,\varpi) = 1$, one has $U(\varpi^\nu, rn) = 0$ whenever $\nu \geqslant \nu(\varpi) + 2$ and $\varpi \neq 3$, and that $U(3^\nu, rn) = 0$ whenever $\nu \geqslant \nu(3) + 3$. Thus we conclude from (7.25) that

$$\mathfrak{A}(\varpi^\nu, n) = 0 \quad \text{for} \quad \nu \geqslant k_0(\varpi, n) + 1, \tag{7.26}$$

where $k_0(\varpi,n) = \nu(\varpi) + 1$ when $\varpi \neq 3$, and $k_0(3,n) = \nu(3) + 2$.

We now continue our investigation of (7.23). As a first step we replace the integral in (7.23) with the integral $J(n)$ defined by (7.8). For each $X > 0$ one readily confirms by means of lemmata 2.7 and 3.3 that

$$\begin{aligned} \int_{XP^{-3}}^{\infty} |v(\beta;P) w(\beta;P)| \, \mathrm{d}\beta &\ll P^4 (\log P)^2 \int_{XP^{-3}}^{\infty} (1 + P^3\beta)^{-5/3} \, \mathrm{d}\beta \\ &\ll P^{1+\varepsilon} \min\{1, X^{-2/3}\}. \end{aligned} \tag{7.27}$$

We take $X = L/q$, and thus deduce from (7.23) that

$$\rho^*(n) - p^{-2} J(n) \sum_{1 \leqslant q \leqslant L} q^{-4} \mathfrak{A}(q,n) \ll P^{1+\varepsilon} L^{-2/3} p^{-2} E_1(n), \tag{7.28}$$

where

$$E_1(n) = \sum_{1 \leqslant q \leqslant L} q^{-10/3} |\mathfrak{A}(q,n)|. \tag{7.29}$$

The next step is to complete the singular series. We note first that in view of our notational conventions, an elementary estimation shows that the series

$$\sum_{q=1}^{\infty} q^{\theta} q_0^3 q_1^2 q_2^{8/3} \tag{7.30}$$

converges whenever $\theta < -3$. Consequently, on using the trivial estimate $(q,n) \leqslant n$ in (7.24), we deduce that for every non-zero integer n the series defined by (7.9) for $\mathfrak{S}(n)$ converges absolutely. Write

$$E_2(n) = \left| \mathfrak{S}(n) - \sum_{1 \leqslant q \leqslant L} q^{-4} \mathfrak{A}(q,n) \right| = \left| \sum_{q > L} q^{-4} \mathfrak{A}(q,n) \right|. \tag{7.31}$$

By (7.8) and (7.27) one has $J(n) \ll P^{1+\varepsilon}$, and thus by (7.28) we deduce that

$$\rho^*(n) - p^{-2} J(n) \mathfrak{S}(n) \ll p^{-2} P^{1+\varepsilon} L^{-2/3} E_1(n) + p^{-2} P^{1+\varepsilon} E_2(n). \tag{7.32}$$

We now estimate $E_1(n)$ and $E_2(n)$ in mean square. An elementary argument provides the estimate

$$\sum_{1 \leqslant n \leqslant N} (q,n)^{1/2} (q',n)^{1/2} \leqslant \mathrm{d}(q) \mathrm{d}(q') N,$$

and hence by (7.24) and the convergence of the series (7.30) for $\theta < -3$, we may conclude from (7.31) that

$$\sum_{1 \leqslant n \leqslant P^3} E_2(n)^2 \ll P^3 \left(\sum_{q > L} q^{\varepsilon - 7/2} q_0^3 q_1^2 q_2^{8/3} \right)^2 \ll P^{3+\varepsilon} L^{-1}. \tag{7.33}$$

A similar argument applied to (7.29) yields the bound

$$\sum_{1 \leqslant n \leqslant P^3} E_1(n)^2 \ll P^{3+\varepsilon} L^{1/3}. \tag{7.34}$$

On substituting (7.33) and (7.34) into (7.32), we reach the conclusion

$$\sum_{1 \leqslant n \leqslant P^3} |\rho^*(n) - p^{-2} J(n) \mathfrak{S}(n)|^2 \ll P^{5+\varepsilon} p^{-4} L^{-1} \ll Q^{5+\varepsilon},$$

so that on recalling (7.22), the proof of the lemma is complete. ■

The proof of theorem 1.2 will be completed by deducing suitable lower bounds for $J(n)$ and $\mathfrak{S}(n)$. In view of (7.3)–(7.5), routine arguments based on Fourier's integral formula (see, for example, Davenport 1959) readily confirm that

$$J(n) \gg P \tag{7.35}$$

for each n satisfying $(1-\tau)N < n \leqslant N$, provided that $\tau > 0$ is sufficiently small. Since the details are standard, we omit them in the interests of saving space. The singular series presents a more challenging problem, and this hurdle we surmount in the following lemma.

Lemma 7.2. *Let X be a real number with $1 \leqslant X \leqslant N$, and let $\mathcal{W}$ denote the set of integers defined in the statement of theorem 1.2. Then the inequality $\mathfrak{S}(n) \geqslant X^{-1}$ holds for all but $O(N^{1+\varepsilon}X^{-1/3})$ integers $n \in \mathcal{W}$ not exceeding N.*

Proof. We begin by rewriting $\mathfrak{S}(n)$ as the Euler product

$$\mathfrak{S}(n) = \prod_{\varpi}(1 + \chi(\varpi, n)), \tag{7.36}$$

where, by (7.9) and (7.26),

$$\chi(\varpi, n) = \sum_{k=1}^{k_0(\varpi,n)} \varpi^{-4k}\mathfrak{A}(\varpi^k, n), \tag{7.37}$$

and k_0 is defined following (7.26). Let $M(q, n)$ denote the number of solutions of the congruence (1.2) with $1 \leqslant x_i, y_i \leqslant q$ $(i = 1, 2)$. Then by a standard argument (see, for example, lemma 2.12 of Vaughan (1997)), one obtains for each K the relation

$$\sum_{k=0}^{K} \varpi^{-4k}\mathfrak{A}(\varpi^k, n) = \varpi^{-3K}M(\varpi^K, n), \tag{7.38}$$

and in particular, in view of (7.37),

$$1 + \chi(\varpi, n) = \varpi^{-3k_0(\varpi,n)}M(\varpi^{3k_0(\varpi,n)}, n). \tag{7.39}$$

It follows that $\mathfrak{S}(n)$ is real and non-negative.

It remains only to bound $\mathfrak{S}(n)$ from below, for $n \in \mathcal{W}$. In order to achieve such a bound, we consider the individual factors in (7.36). Suppose first that $\varpi \nmid 3n$. Then $k_0(\varpi, n) = 1$, and (7.39) yields $1 + \chi(\varpi, n) \geqslant \varpi^{-3}$ for $n \in \mathcal{W}$. By (7.37) and (7.24), moreover, for primes ϖ of the latter type one has $|\chi(\varpi, n)| \ll \varpi^{-3/2}$. Thus we deduce that

$$\prod_{\varpi \nmid 3n} (1 + \chi(\varpi, n)) \gg 1,$$

where the implicit constant depends at most on D_1 and D_2. Now suppose that $\varpi | n$, but $\varpi \nmid 6D_1D_2$. In this case we deduce from (7.10), (7.24), (7.37), together with lemmata 2.4 and 2.5, that for some positive constant C one has

$$|\chi(\varpi, n)| \leqslant \frac{81}{\varpi} + \frac{1}{\varpi^2} + \frac{1}{2}C\sum_{k=3}^{\infty} \varpi^{-k/3} \leqslant \frac{C}{\varpi}.$$

We may suppose without loss of generality that $C > 6\max\{|D_1|, |D_2|\}$, and then deduce that

$$\prod_{\substack{\varpi | n \\ \varpi > C}} (1 + \chi(\varpi, n)) \geqslant \prod_{\substack{\varpi | n \\ \varpi > C}} (1 - C/\varpi) \gg (\log\log n)^{-C}. \tag{7.40}$$

Continuing to work with this constant C, we define now the function

$$s(n) = \prod_{\substack{\varpi^\nu \| n \\ \varpi \leqslant C}} \varpi^\nu. \tag{7.41}$$

When $n \in \mathcal{W}$, it follows from (7.39) that

$$\prod_{\substack{\varpi \leqslant C \\ \varpi | 3n}} (1 + \chi(\varpi, n)) \geqslant \prod_{\varpi \leqslant C} \varpi^{-3k_0(\varpi, n)} \geqslant (3s(n))^{-3} \prod_{\varpi \leqslant C} \varpi^{-3} \gg s(n)^{-3}. \tag{7.42}$$

Combining the conclusions (7.40) and (7.42), we discover that for each $n \in \mathcal{W}$ one has

$$\mathfrak{S}(n) \gg (\log n)^{-1} s(n)^{-3}, \tag{7.43}$$

where the implicit constant depends at most on the discriminants D_1 and D_2.

Next let Z satisfy $1 \leqslant Z \leqslant N$, and consider

$$\mathcal{Z} = \operatorname{card}\{1 \leqslant n \leqslant N : s(n) > Z\}. \tag{7.44}$$

Let $\mathcal{S}$ denote the set of integers not exceeding N all of whose prime factors are at most C in size. Then plainly $\operatorname{card}(\mathcal{S}) \ll (\log N)^C$, and thus, on recalling (7.41) we have

$$\begin{aligned} \mathcal{Z} &\leqslant \sum_{\substack{s \in \mathcal{S} \\ s > Z}} \operatorname{card}\{1 \leqslant n \leqslant N : s(n) = s\} \\ &\leqslant \sum_{\substack{s \in \mathcal{S} \\ s > Z}} \operatorname{card}\{1 \leqslant n \leqslant N : s | n\} \ll \frac{N}{Z} (\log N)^C. \end{aligned} \tag{7.45}$$

To complete the proof of the lemma we have only to collect together (7.43), (7.44) and (7.45), and put $Z = B(X/\log N)^{1/3}$, for a suitable positive constant B.

Finally, we prove that $\mathcal{W}$ has positive density by constructing a fixed arithmetic progression contained in $\mathcal{W}$, thereby establishing the first claim of theorem 1.2.

Lemma 7.3. *The set $\mathcal{W}$ contains an arithmetic progression with modulus depending only on the discriminants D_1 and D_2.*

Proof. We write

$$\Phi_j(x, y) = a_j x^3 + b_j x^2 y + c_j x y^2 + d_j y^3 \quad (j = 1, 2),$$

and denote the highest common factor of the eight coefficients a_j, b_j, c_j, d_j $(j = 1, 2)$ by K. We may suppose that $K = 1$, for otherwise we may consider the integers represented in the form

$$K^{-1}(\Phi_1(x_1, y_1) + \Phi_2(x_2, y_2)), \tag{7.46}$$

and if it were known that the congruence conditions were satisfied for the form (7.46) for a positive proportion of the integers, then of course the same is true for $\Phi_1(x_1, y_1) + \Phi_2(x_2, y_2)$.

We consider next a fixed prime ϖ, and aim to construct a non-singular solution of the congruence

$$\Phi_1(x_1, y_1) + \Phi_2(x_2, y_2) \equiv n \pmod{\varpi}, \tag{7.47}$$

that is, a solution of (7.47) for which at least one of the partial derivatives

$$\frac{\partial \Phi_i}{\partial x_i} \quad \text{and} \quad \frac{\partial \Phi_i}{\partial y_i} \quad (i = 1, 2),$$

does not vanish.

Suppose first that $\varpi \nmid 6D_1D_2$. Then by (7.10) and lemma 2.4, one has

$$|\mathfrak{A}(\varpi, n)| \leqslant 81(\varpi - 1)\varpi^2,$$

and so (7.38) shows that whenever $\varpi > 100$, one has

$$M(\varpi, n) \geqslant \varpi^3 - 81\varpi(\varpi - 1) \geqslant \tfrac{1}{6}\varpi^3. \tag{7.48}$$

However, the solutions of the congruence (7.47) singular modulo ϖ must also satisfy the simultaneous congruences

$$3a_jx_j^2 + 2b_jx_jy_j + c_jy_j^2 \equiv b_jx_j^2 + 2c_jx_jy_j + 3d_jy_j^2 \equiv 0 \pmod{\varpi} \quad (j = 1, 2). \tag{7.49}$$

But if ϖ does not divide any of the coefficients in (7.49), then the congruences (7.49) possess at most $O(\varpi)$ solutions with $1 \leqslant x_j, y_j \leqslant \varpi$, for each of $j = 1$ and 2. It follows that for any integer n, the total number of solutions to the congruence (7.47) singular modulo ϖ is at most $O(\varpi^2)$, whence by (7.48) there exists a positive constant C such that, for all $\varpi > C$, the congruence (7.47) possesses a solution non-singular modulo ϖ. By a standard application of Hensel's lemma, therefore, we deduce that for each $\varpi > C$, and for all natural numbers k and integers n, one has

$$M(\varpi^k, n) \geqslant \varpi^{3(k-1)}. \tag{7.50}$$

Consider next the primes ϖ with $3 < \varpi \leqslant C$. On recalling that $K = 1$, we have that ϖ does not divide all of the coefficients of the four polynomials in (7.49), and therefore at least one of these polynomials does not vanish identically modulo ϖ. We may therefore pick integers $(x_j(\varpi), y_j(\varpi))$ $(j = 1, 2)$ for which at least one of the four congruences (7.49) fails. Write

$$n_0(\varpi) = \Phi_1(x_1(\varpi), y_1(\varpi)) + \Phi_2(x_2(\varpi), y_2(\varpi)).$$

Then all integers $n \equiv n_0(\varpi) \pmod{\varpi}$ admit a non-singular solution to (7.47), by construction. Thus, for all $k \geqslant 1$ and each $n \equiv n_0(\varpi) \pmod{\varpi}$, Hensel's lemma again yields $M(\varpi^k, n) \geqslant 1$. When $\varpi = 2$ or $\varpi = 3$, the above argument is readily modified by considering congruences modulo 8 and 27 respectively, such moduli being required to lift solutions of the cubic congruence to higher powers of 2 and 3, respectively. On recalling (7.50), it follows that the arithmetic progression defined by $n \equiv n_0(\varpi) \pmod{\varpi}$ $(3 < \varpi \leqslant C)$, and $n \equiv n_0(8) \pmod 8$, $n \equiv n_0(27) \pmod{27}$, is contained in $\mathcal{W}$. This completes the proof of the lemma. ■

Our preparations complete, we now deliver the coup de grace by popping theorem 1.2 into the back of the net. By lemma 7.2 and (7.35), we have

$$p^{-2}J(n)\mathfrak{S}(n) \gg p^{-2}PX^{-1} \tag{7.51}$$

for all but $O(N^{1+\varepsilon}X^{-1/3})$ of the integers $n \in \mathcal{W}$ with $(1 - \tau)N < n \leqslant N$. By lemma 7.1, the inequality

$$|\rho(n) - p^{-2}J(n)\mathfrak{S}(n)| > p^{-2}PX^{-1}(\log P)^{-1}$$

can hold for at most $\mathcal{E}$ natural numbers n with $1 \leqslant n \leqslant N$, where

$$\mathcal{E} \ll Q^{5+\varepsilon}(p^{-2}P^{1-\varepsilon}X^{-1})^{-2} \ll N^{1+\varepsilon}X^2p^{-1} \ll N^{1-(1/30)+\varepsilon}X^2. \tag{7.52}$$

We choose $X = N^{1/70}$, and conclude from (7.51) and (7.52) that $\rho(n) \gg p^{-2}P^{67/70}$ for all $n \in \mathcal{W}$ with $(1 - \tau)N < n \leqslant N$, with the exception of at most $O(N^{209/210+\varepsilon})$ integers. The claimed bound on the exceptional set recorded in theorem 1.2 then follows by a standard argument, dividing up the interval $[1, N]$ into subintervals of the shape $[(1-\tau)N', N']$. We have thus scored our final goal of establishing theorem 1.2.

T.D.W. is a Packard Fellow, and supported in part by NSF grant DMS-9622773.

References

Baker, R. C. 1989 Diagonal cubic equations. II. *Acta Arith.* **53**, 217–250.

Brüdern, J. 1991 Ternary additive problems of Waring's type. *Math. Scand.* **68**, 27–45.

Chowla, S. & Davenport, H. 1961 On Weyl's inequality and Waring's problem for cubes. *Acta Arith.* **6**, 505–521.

Davenport, H. 1939 On Waring's problem for cubes. *Acta Math.* **71**, 123–143.

Davenport, H. 1959 Cubic forms in thirty-two variables. *Phil. Trans. R. Soc. Lond.* A **251**, 193–232.

Davenport, H. 1962 Cubic forms in 29 variables. *Proc. R. Soc. Lond.* A **266**, 287–298.

Davenport, H. 1963 Cubic forms in sixteen variables. *Proc. R. Soc. Lond.* A **272**, 285–303.

Davenport, H. & Heilbronn, H. 1937 On Waring's problem: two cubes and one square. *Proc. Lond. Math. Soc.* (2) **43**, 73–104.

Heath-Brown, D. R. 1983 Cubic forms in ten variables. *Proc. Lond. Math. Soc.* (3) **47**, 225–257.

Hooley, C. 1967 On binary cubic forms. *J. Reine Angew. Math.* **226**, 30–87.

Hooley, C. 1985 On the representations of numbers by binary cubic forms. *Glasgow Math. J.* **27**, 95–98.

Hooley, C. 1986 On Waring's problem. *Acta Math.* **157**, 49–97.

Hooley, C. 1988 On nonary cubic forms. *J. Reine Angew. Math.* **386**, 32–98.

Hooley, C. 1991 On nonary cubic forms. II. *J. Reine Angew. Math.* **415**, 95–165.

Hooley, C. 1994 On nonary cubic forms. III. *J. Reine Angew. Math.* **456**, 53–63.

Hua, L.-K. 1938 On the representation of numbers as the sums of the powers of primes. *Math. Z.* **44**, 335–346.

Lewis, D. J. 1957*a* Cubic forms over algebraic number fields. *Mathematika* **4**, 97–101.

Lewis, D. J. 1957*b* Cubic congruences. *Michigan Math. J.* **4**, 85–95.

Schmidt, W. M. 1976 *Equations over finite fields. An elementary approach. Springer Lecture Notes in Mathematics*, vol. 365. Berlin: Springer.

Vaughan, R. C. 1986 On Waring's problem for cubes. *J. Reine Angew. Math.* **365**, 122–170.

Vaughan, R. C. 1989 On Waring's problem for cubes. II. *J. Lond. Math. Soc.* (2) **39**, 205–218.

Vaughan, R. C. 1997 *The Hardy–Littlewood method*, Cambridge tract no. 125, 2nd edn. Cambridge University Press.

Vaughan, R. C. & Wooley, T. D. 1991 Waring's problem: some refinements. *Proc. Lond. Math. Soc.* (3) **63**, 35–68.

Vaughan, R. C. & Wooley, T. D. 1994 Further improvements in Waring's problem, II: Sixth powers. *Duke Math. J.* **76**, 683–710.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES