Linear relations amongst sums of two squares

D.R. Heath-Brown

1 Introduction

It is well known that there are infinitely many sets of three distinct primes in arithmetic progression. This may be proved by an easy adaptation of Vinogradov's treatment of the ternary Goldbach problem. More generally for, any nonzero integers A, B, C, not all of the same sign, one can show the existence of infinitely many triples of primes p_1, p_2, p_3 satisfying the linear relation

$$Ap_1 + Bp_2 + Cp_3 = 0$$

subject to the natural condition that $A+B+C$ should be even. Balog [1] has made important progress on the question of linear relations involving more than 3 primes, but nonetheless it remains an open problem as to whether there are infinitely many sets of 4 distinct primes in arithmetic progression.

Many open problems involving primes have potentially easier relatives involving sums of two squares. Thus one might ask whether or not there are infinitely many arithmetic progressions of 4 (or more) distinct integers, each of which is a sum of 2 squares. This is trivial. The numbers

$$(n-8)^2 + (n-1)^2, \quad (n-7)^2 + (n+4)^2, \quad (n+7)^2 + (n-4)^2$$
$$\text{and} \quad (n+8)^2 + (n+1)^2$$

form an arithmetic progression with common difference $12n$. In this paper we shall address the question of the frequency of such progressions. We shall count the sums of two squares with appropriate multiplicity, so that we shall consider the sum

$$\sum_{\mathbf{x} \in \mathcal{R}} r(L_1(\mathbf{x}))r(L_2(\mathbf{x}))r(L_3(\mathbf{x}))r(L_4(\mathbf{x})), \tag{1.1}$$

where $\mathcal{R}$ is a suitable subset of $\mathbb{R}^2$ and the linear forms L_i are given by

$$\begin{array}{ll} L_1(\mathbf{x}) = x_1, & L_2(\mathbf{x}) = x_1 + x_2, \\ L_3(\mathbf{x}) = x_1 + 2x_2, & L_4(\mathbf{x}) = x_1 + 3x_2, \end{array} \tag{1.2}$$

where $\mathbf{x}$ denotes the vector (x_1, x_2). The corresponding problem for arithmetic progressions of length 3 is readily handled by the circle method. However for progressions of length 4 it would appear that one would require a version of the 'Kloosterman refinement' for a double integral

$$\int_0^1 \int_0^1 S(\alpha)^2 S(-2\alpha+\beta)^2 S(\alpha-2\beta)^2 S(\beta)^2 d\alpha d\beta.$$

Since research to date has failed to provide such a technique we shall use a rather different approach.

We shall consider a general set of linear forms $L_1, \ldots, L_4$. However we will find it convenient to work with linear forms which are suitably normalized. Moreover we shall require the region $\mathcal{R}$ in which we work to satisfy certain basic conditions. We therefore introduce the following hypothesis.

Normalization Condition 1 (NC1) *We assume:*

(i) No two of the forms $L_1, \ldots, L_4$ are proportional.

(ii) We have

$$\mathcal{R} = X\mathcal{R}^{(0)} = \{\mathbf{x} \in \mathbb{R}^2 : X^{-1}\mathbf{x} \in \mathcal{R}^{(0)}\},$$

where $\mathcal{R}^{(0)} \subset \mathbb{R}^2$ is open, bounded and convex, with a piecewise continuously differentiable boundary, and where X is a large positive parameter.

(iii) We have $L_i(\mathbf{x}) > 0$ for $1 \le i \le 4$ and for all $\mathbf{x} \in \mathcal{R}^{(0)}$.

(iv) We have

$$L_1(x_1, x_2) \equiv L_2(x_1, x_2) \equiv L_3(x_1, x_2) \equiv L_4(x_1, x_2) \equiv x_1 \pmod 4.$$

We have imposed the final condition in order to simplify our analysis. While this may seem a little arbitrary, it can be viewed as an analogue of conditions (ii) and (iii). One can think of (ii) and (iii) as requiring $\mathbf{x}$ to lie in an open neighbourhood of a point $\mathbf{y}$ for which each $L_i(\mathbf{y})$ is a sum of two squares. The 2-adic analogue of this real condition on the domain of summation would involve fixing a 2-adic vector $\mathbf{y}$ such that each value $L_i(\mathbf{y})$ is a sum of two 2-adic squares. We would then require $\mathbf{x}$ to lie in an appropriate 2-adic neighbourhood of $\mathbf{y}$. If one imposes such a condition then it can be shown that there is a suitable change of variables which produces forms satisfying (iv). However we shall not pursue this here.

In view of condition (iv) we shall find it convenient to write

$$\mathcal{R}_4 = \{\mathbf{x} \in \mathcal{R} : x_1 \equiv 1 \pmod 4\},$$

so that our problem is to estimate

$$\sum_{\mathbf{x}\in\mathcal{R}_4} r(L_1(\mathbf{x}))r(L_2(\mathbf{x}))r(L_3(\mathbf{x}))r(L_4(\mathbf{x})) = S, \tag{1.3}$$

say.

From now on, all order constants will be allowed to depend on the set of forms $L_1, \ldots, L_4$, and on the region $\mathcal{R}^{(0)}$. Our first result is then the following.

Theorem 1 *For a set of forms satisfying* **NC1**, *we have*

$$S = 4\pi^4 \operatorname{meas}\mathcal{R}\prod_{p\geq 3}\sigma_p + O(X^2(\log X)^{-\eta/2}(\log\log X)^{15/4}) \tag{1.4}$$

where meas *denotes Lebesgue measure, and*

$$\eta = 1 - \frac{1+\log\log 2}{\log 2} = 0.08607\ldots. \tag{1.5}$$

Here the product $\prod \sigma_p$ is absolutely convergent and

$$\sigma_p = E_p\{1-\chi(p)p^{-1}\}^4,$$

where χ is the nonprincipal character modulo 4. The factor E_p is given by

$$E_p = \sum_{a,b,c,d=0}^{\infty} \chi(p)^{a+b+c+d}\rho(p^a,p^b,p^c,p^d)^{-1},$$

where $\rho(d_1,d_2,d_3,d_4)$ is the determinant of the lattice

$$\big\{\mathbf{x}\in\mathbb{Z}^2 : d_i \mid L_i(\mathbf{x}),\ 1\leq i\leq 4\big\}.$$

The implied constant in (1.4) may depend on the set of forms $L_1, \ldots, L_4$, and on the region $\mathcal{R}^{(0)}$.

It may be of interest to note that we can evaluate E_p explicitly in many cases. For $1 \leq i < j \leq 4$, let Δ_{ij} be the determinant of the pair of forms L_i, L_j, and let Δ be the product of the various Δ_{ij}. Then if $p \nmid \Delta$, we can find E_p by a routine, if lengthy, calculation. The result is that

$$E_p = \begin{cases} (1-\frac{1}{p})^{-2}(1-\frac{1}{p^2})^{-2}(1+\frac{2}{p}+\frac{6}{p^2}+\frac{2}{p^3}+\frac{1}{p^4}) & \text{if } \chi(p)=1, \\ (1-\frac{1}{p^2})^{-1}(1-\frac{1}{p^4})^{-1}(1-\frac{1}{p})^4 & \text{if } \chi(p)=-1. \end{cases} \tag{1.6}$$

It follows in particular that $\prod \sigma_p = 0$ if and only if there is some prime $p \mid \Delta$ with $\chi(p) = -1$ for which $E_p = 0$.

It is perhaps worth observing that a notional application of the Hardy–Littlewood circle method to the system

$$L_i(x_1, x_2) = u_i^2 + v_i^2, \quad (1 \le i \le 4),$$

consisting of 4 equations in 10 variables predicts exactly the main term given in (1.4). In particular, the singular integral (the density for the real valuation) is $\pi^4 \operatorname{meas} \mathcal{R}$, and the 2-adic density

$$\lim_{n\to\infty} \#\{\mathbf{x}, \mathbf{u}, \mathbf{v} \pmod{2^n} : x_1 \equiv 1 \pmod 4,\ L_i(\mathbf{x}) \equiv u_i^2 + v_i^2 \pmod{2^n}\}$$

is 4.

To apply Theorem 1 to arithmetic progressions of length 4 we note that if 4 integers in arithmetic progression are each a sum of two squares, then the common difference must be a multiple of 4. Take

$$\mathcal{R} = \{(x_1, x_2) \in \mathbb{R}^2 : x_1, x_2 > 0,\ x_1 + 12x_2 < X\}$$

and

$$L_1(\mathbf{x}) = x_1, \quad L_2(\mathbf{x}) = x_1 + 4x_2, \quad L_3(\mathbf{x}) = x_1 + 8x_2, \quad L_4(\mathbf{x}) = x_1 + 12x_2.$$

Since $r(2n) = r(n)$ we see that

$$\sum_{a<b<c<d<X} r(a)r(b)r(c)r(d) = \sum_k \sum_{\substack{2^k(x_1,x_2)\in\mathcal{R} \\ 2\nmid x_1}} r(L_1(\mathbf{x}))r(L_2(\mathbf{x}))r(L_3(\mathbf{x}))r(L_4(\mathbf{x})),$$

where the sum over a, b, c, d is restricted to arithmetic progressions of length 4. Now if we set

$$\mathcal{R}_4(k) = \{(x_1, x_2) \in \mathbb{Z}^2 : 2^k(x_1, x_2) \in \mathcal{R},\ x_1 \equiv 1 \pmod 4\},$$

we see that

$$\sum_{a<b<c<d<X} r(a)r(b)r(c)r(d) = \sum_k \sum_{(x_1,x_2)\in\mathcal{R}_4(k)} r(L_1(\mathbf{x}))r(L_2(\mathbf{x}))r(L_3(\mathbf{x}))r(L_4(\mathbf{x})).$$

We have sufficient uniformity in Theorem 1 to sum over k. Since $\operatorname{meas} \mathcal{R} = X^2/24$ and $\sum_0^\infty 4^{-k} = 4/3$, this therefore yields the asymptotic formula

$$\sum_{a<b<c<d<X} r(a)r(b)r(c)r(d) = CX^2 + O(X^2(\log X)^{-\eta/2}(\log\log X)^{15/4}),$$

Table 1

X	$S(X)$	$S(X)/CX^2$
1000	21833216	$21.833\ldots$
2000	91315200	$22.828\ldots$
4000	381608960	$23.850\ldots$
8000	1554144256	$24.283\ldots$
16000	6308194304	$24.641\ldots$
32000	25428982272	$24.832\ldots$
64000	102495412736	$25.023\ldots$
128000	411816625664	$25.135\ldots$

where the sum over a, b, c, d is restricted to arithmetic progressions of length 4. Since $\operatorname{meas}\mathcal{R} = X^2/24$, the constant C takes the form

$$C = 4\pi^4 \frac{1}{24}\frac{4}{3}\prod_{p\geq 3} E_p\{1-\chi(p)p^{-1}\}^4,$$

with E_p given by (1.6) for $p \geq 5$. Moreover one may compute that

$$E_3 = \frac{27}{80}.$$

Since progressions with $d = X$ clearly contribute $O(X^{1+\varepsilon})$ for any $\varepsilon > 0$ we may summarize our conclusion as follows.

Corollary 1 *There is a positive constant C such that*

$$\sum_{a<b<c<d\leq X} r(a)r(b)r(c)r(d) = CX^2 + O(X^2(\log X)^{-\eta/2}(\log\log X)^{15/4}),$$

where the sum over a, b, c, d is restricted to arithmetic progressions of length 4. The constant C has the approximate value $25.3039\ldots$.

The corollary is illustrated by Table 1, in which

$$S(X) = \sum_{a<b<c<d<X} r(a)r(b)r(c)r(d).$$

The general problem as formulated above is relevant to a very different question. The simultaneous equations

$$V: \quad \begin{cases} L_1(x_1,x_2)L_2(x_1,x_2) = x_3^2 + x_4^2 \\ L_3(x_1,x_2)L_4(x_1,x_2) = x_5^2 + x_6^2 \end{cases} \tag{1.7}$$

will, in general, define a 3-fold in $\mathbb{P}^5$. We can estimate the number of rational points on this variety as $\mathbf{x}$ runs over a region $\mathcal{R}$ by examining the sum

$$\sum_{\mathbf{x}\in\mathcal{R}} r(L_1(\mathbf{x})L_2(\mathbf{x}))r(L_3(\mathbf{x})L_4(\mathbf{x})).$$

Varieties of the type (1.7) are of considerable interest, since they may fail to satisfy the Hasse Principle. Thus they may have no nontrivial rational points even though they have nonsingular points over $\mathbb{R}$ and each of the p-adic fields $\mathbb{Q}_p$. For general pairs of quadratic forms this observation is due to Iskovskih [6]. For varieties of the particular shape (1.7) the phenomenon is illustrated by the example

$$x_1x_2 = x_3^2 + x_4^2, \qquad (3x_1 + 4x_2)(8x_1 + 11x_2) = x_5^2 + x_6^2, \tag{1.8}$$

as we proceed to show. There are nonsingular points with $x_1 = x_2 = 1$ in $\mathbb{R}$ and in $\mathbb{Q}_p$ for every prime p other than $p = 7$ and $p = 19$. Similarly for these two exceptional fields there are nonsingular points with $x_1 = 2$ and $x_2 = 1$. We proceed to assume that the equations (1.7) have a nonzero integral solution $x_1, \ldots, x_6$. In particular it follows that x_1 and x_2 cannot both be zero. For any $d \in \mathbb{N}$, if nd^2 is a sum of two squares, then n is also a sum of two squares. Thus we may assume, without loss of generality, that x_1 and x_2 are coprime. Moreover, we may change the signs if necessary, so as to suppose that at least one of x_1 and x_2 is positive. Then, since their product is a sum of two squares, we see that the other must be nonnegative. It follows firstly that each of x_1 and x_2 is a sum of two squares, and secondly that each of $3x_1 + 4x_2$ and $8x_1 + 11x_2$ is strictly positive. Now

$$\begin{vmatrix} 3 & 4 \\ 8 & 11 \end{vmatrix} = 1,$$

so that $3x_1 + 4x_2$ and $8x_1 + 11x_2$ must be coprime. Thus both $3x_1 + 4x_2$ and $8x_1 + 11x_2$ will be sums of two squares.

Now if x_1 is odd, then $x_1 = a^2 + b^2 \equiv 1 \pmod 4$, so that we must have $3x_1 + 4x_2 \equiv 3 \pmod 4$. Thus $3x_1 + 4x_2$ cannot be a sum of two squares. Similarly if x_1 is even, then x_2 must be odd, and hence $x_2 \equiv 1 \pmod 4$, since x_2 is a sum of two squares. However this means that $8x_1 + 11x_2 \equiv 3 \pmod 4$ so that $8x_1 + 11x_2$ cannot be a sum of two squares. This completes the proof.

Even when the variety does possess rational points, it may fail to satisfy the weak approximation principle. In general, a variety V is said to satisfy the *weak approximation principle* if its rational points are dense in the adélic points. To put this in concrete terms, for our variety (1.7), suppose we are given a real point $(x_1^{(\mathbb{R})}, \ldots, x_6^{(\mathbb{R})})$ and p-adic points $(x_1^{(p)}, \ldots, x_6^{(p)})$ for a finite number of distinct primes p, all lying on the variety (1.7). The weak approximation principle then asserts that, for any $\varepsilon > 0$, we can find a rational point $(x_1, \ldots, x_6)$ on (1.7) satisfying the simultaneous conditions

$$|x_i - x_i^{(\mathbb{R})}| < \varepsilon \quad \text{and} \quad |x_i - x_i^{(p)}|_p < \varepsilon, \quad (1 \le i \le 6)$$

for each of the primes p.

However it can happen that V fails to satisfy even the real condition. In particular the variety may have two real components, on one of which the rational points are dense, and on the other of which there are no rational points. This is demonstrated by the example

$$x_1x_2 = x_3^2 + x_4^2, \qquad (x_1 - x_2)(3x_1 - 8x_2) = x_5^2 + x_6^2, \tag{1.9}$$

due to Colliot-Thélène, Coray and Sansuc [2]. There is clearly a rational point with $x_1 = 1$ and $x_2 = 2$. Moreover the real points belong to two components, namely those with $x_2/x_1 \ge 1$ and $0 \le x_2/x_1 \le 3/8$. (We regard points with $x_1 = 0$ as being of the first type.) The special feature of this example is that all rational points lie on the first of these components. To prove this we shall suppose we have an integer point for which $0 \le x_2/x_1 \le 3/8$, and derive a contradiction. As with (1.8) we may assume that x_1 and x_2 are coprime and nonnegative, so that they must both be sums of two squares. Our assumption on the size of x_2/x_1 implies that $x_1 - x_2$ and $3x_1 - 8x_2$ are both nonnegative. Since

$$\begin{vmatrix} 1 & -1 \\ 3 & -8 \end{vmatrix} = -5,$$

the highest common factor of $x_1 - x_2$ and $3x_1 - 8x_2$ must be either 1 or 5. Thus, since the product of the linear forms $x_1 - x_2$ and $3x_1 - 8x_2$ is a sum of two squares, they must each be a sum of two squares.

Now if x_1 is odd, then $x_1 = a^2 + b^2 \equiv 1 \pmod 4$, so that we must have $3x_1 - 8x_2 \equiv 3 \pmod 4$. Thus $3x_1 - 8x_2$ cannot be a sum of two squares. Similarly if $2 \parallel x_1$ we will have $x_1 \equiv 2 \pmod 8$ and $3x_1 - 8x_2 \equiv 6 \pmod 8$, so that $3x_1 - 8x_2$ is not a sum of two squares. Finally, if $4 \mid x_1$, then x_2 is odd, and we will have $x_2 = c^2 + d^2 \equiv 1 \pmod 4$. In this case $x_1 - x_2 \equiv 3 \pmod 4$ and $x_1 - x_2$ cannot be a sum of two squares. This establishes our claim.

In general there is a heuristic expectation that the number of rational points on a given variety which lie in a large region should be given by a product of local densities. This is indeed the type of asymptotic formula that the Hardy–Littlewood circle method provides, in those cases for which the error terms can be successfully estimated. However when the rational points on a variety are not evenly distributed amongst the admissible adélic points, the entire rationale for this heuristic expectation breaks down. It is thus of considerable interest to estimate the number of points on such a variety, and to compare the result with that predicted from the product of local densities. This is what we shall do for the varieties (1.7).

We shall introduce the same type of normalization condition as before. Specifically, we require the following:

Normalization Condition 2 (NC2) *We assume:*

(i) No two of the forms $L_1, \ldots, L_4$ are proportional.

(ii) We have

$$\mathcal{R} = X\mathcal{R}^{(0)} = \{\mathbf{x} \in \mathbb{R}^2 : X^{-1}\mathbf{x} \in \mathcal{R}^{(0)}\},$$

where $\mathcal{R}^{(0)} \subset \mathbb{R}^2$ is open, bounded and convex, with a piecewise continuously differentiable boundary, and where X is a large positive parameter.

(iii) We have $L_i(\mathbf{x}) > 0$ for $1 \leq i \leq 4$ and for all $\mathbf{x} \in \mathcal{R}^{(0)}$.

(iv) We have

$$L_1(x_1, x_2) \equiv L_2(x_1, x_2) \equiv \nu x_1 \pmod 4$$

and

$$L_3(x_1, x_2) \equiv L_4(x_1, x_2) \equiv \nu' x_1 \pmod 4,$$

for appropriate $\nu, \nu' = \pm 1$.

In connection with condition (iii) we note that the equations (1.7) do not require that $L_i(\mathbf{x}) > 0$. However, apart from $O(X)$ points where some L_i vanishes, the solutions may be subdivided into regions in which each L_i is one signed. On each such region we can then replace L_i by $\pm L_i$ as necessary, so as to ensure that we have points with $L_i(\mathbf{x}) > 0$.

As with **NC1**, condition (iv) is imposed in order to simplify the exposition. However it may be viewed, as before, as being the result of restricting $\mathbf{x}$ to a suitable 2-adic region.

As an example, we note that the variety defined by (1.8) has a 2-adic point $x_1^{(0)}, \ldots, x_6^{(0)}$ with $x_1^{(0)} = x_2^{(0)} = 1$. The region given by $x_1 - x_2 \equiv x_1^{(0)} - x_2^{(0)} \equiv 0 \pmod 4$ is a 2-adic neighbourhood of the point $x_1^{(0)}, \ldots, x_6^{(0)}$. For any point in this neighbourhood we may write $x_1 = y_1$ and $x_2 = y_1 + 4y_2$ to produce the equations

$$y_1(y_1 + 4y_2) = x_3^2 + x_4^2, \qquad (7y_1 + 16y_2)(19y_1 + 44y_2) = x_5^2 + x_6^2. \tag{1.10}$$

The linear forms now satisfy part (iv) of **NC2**.

Similarly for the example (1.9) we have a 2-adic point with $x_1^{(0)} = 1$ and $x_2^{(0)} = 2$, and we use the 2-adic region

$$x_2 - 2x_1 \equiv x_2^{(0)} - 2x_1^{(0)} \equiv 0 \pmod{8}.$$

We thus write $x_1 = y_1$ and $x_2 = 2y_1 + 8y_2$ to produce the equations

$$y_1(y_1 + 4y_2) = y_3^2 + y_4^2, \qquad (y_1 + 8y_2)(13y_1 + 64y_2) = x_5^2 + x_6^2, \tag{1.11}$$

all of whose rational points we have shown to satisfy $y_2/y_1 \geq -1/8$. Again the linear forms satisfy part (iv) of **NC2**.

In view of part (iv) of **NC2** it is natural to restrict consideration to the case in which (x_1, x_2) lies in the set

$$\mathcal{R}_2 = \{\mathbf{x} \in \mathcal{R} : x_1 \equiv 1 \pmod{2}\}.$$

Our principal result describing the number of rational points on the general variety (1.7) is now as follows.

Theorem 2 *Suppose* **NC2** *holds. The local densities for the variety V with equations (1.7), for the set $\mathcal{R}_2$, are given by*

$$\sigma_\infty = \pi^2 \operatorname{meas} \mathcal{R}, \qquad \sigma_2 = 2$$

and

$$\sigma_p = (1 - \chi(p)/p)^2 T_\chi(p), \qquad (p \geq 3), \tag{1.12}$$

where

$$T_\chi(p) = E_p^{(0,0)} - \chi(p)E_p^{(0,1)} - \chi(p)E_p^{(1,0)} + E_p^{(1,1)} \tag{1.13}$$

and

$$E_p^{(u,v)} = \sum_{\alpha,\beta,\gamma,\delta=0}^{\infty} \chi(p)^{\alpha+\beta+\gamma+\delta} \rho(p^{\alpha+u}, p^{\beta+u}, p^{\gamma+v}, p^{\delta+v})^{-1}. \tag{1.14}$$

Here $\rho(d_1, d_2, d_3, d_4)$ is as in Theorem 1. Moreover, when $p \nmid \Delta$ we have

$$\sigma_p = (1 + \chi(p)/p)^2. \tag{1.15}$$

If $\sigma_p = 0$ for any prime p then V has no rational point with $(x_1, x_2) \in \mathcal{R}_2$. If $\sigma_p \neq 0$ for every prime p, then

$$\sum_{\mathbf{x} \in \mathcal{R}_2} r(L_1(\mathbf{x})L_2(\mathbf{x}))r(L_3(\mathbf{x})L_4(\mathbf{x})) = \{1 + \varepsilon\}\sigma_\infty \prod_p \sigma_p + o(X^2).$$

where

$$\varepsilon = \chi(\nu\nu') \prod_{p|\Delta,\, \chi(p)=-1} T_-(p)/T_+(p), \tag{1.16}$$

with

$$T_\pm(p) = E_p^{(0,0)} \pm E_p^{(0,1)} \pm E_p^{(1,0)} + E_p^{(1,1)}. \tag{1.17}$$

Moreover, when $p \equiv -1 \pmod 4$ *we have* $E_p^{(u,v)} \geq 0$*, so that*

$$|T_-(p)| \leq T_+(p).$$

We also have $E_p^{(1,0)} = E_p^{(0,1)} = 0$ *for any prime* $p \equiv -1 \pmod 4$ *not dividing* $\Delta_{12}\Delta_{34}$*.*

If $\varepsilon = -1$ *then* V *has no rational point with* $(x_1, x_2) \in \mathcal{R}_2$*.*

Thus the factor $1+\varepsilon$ measures the discrepancy between the true asymptotic formula and the Hardy–Littlewood prediction. Although we shall not prove it here, we may remark that the sums $T_\pm(p)$ are always rational numbers, so that the factor $1+\varepsilon$ is a rational number in the range $[0,2]$.

We see that Theorem 2 establishes a local to global principle in the shape of the assertion that if $\sigma_p > 0$ for every p, then there exist rational points on V, providing that $1+\varepsilon \neq 0$. Moreover it is a standard fact that we will have $\sigma_p > 0$ for any prime for which V has a nonsingular p-adic point. In contrast, our result does not give a full solution to the weak approximation problem, since we are unable to restrict the variables x_3, x_4, x_5, x_6 in (1.7). However, we are able to control the variables x_1, x_2 by our method.

In fact it is known that the Brauer–Manin obstruction is the only obstruction to both the Hasse Principle and Weak Approximation, for varieties of the form (1.7). Although this is not formally stated in the literature, it is possible to use a descent argument to reduce the problem to one involving a certain intersection of two quadrics in $\mathbb{P}^6$, to which Theorem 6.7 of Colliot-Thélène, Sansuc and Swinnerton-Dyer [3] may be applied. In particular it follows that our condition $1+\varepsilon > 0$ must be equivalent to the emptiness of the Brauer–Manin obstruction for the Hasse Principle.

In the final section of the paper we shall investigate the examples (1.8) and (1.9) more fully, as well as the variety

$$x_1(x_1 + 12x_2) = x_3^2 + x_4^2, \qquad (x_1 + 4x_2)(x_1 + 16x_2) = x_5^2 + x_6^2, \tag{1.18}$$

for which we shall show that $0 < 1+\varepsilon < 2$.

We conclude this introduction by remarking that it should be possible to replace the character χ by any other nonprincipal real character. Indeed one should be able to use different characters for each of the four linear forms in Theorem 1. In the same way, in Theorem 2 one would take any two nonprincipal real characters χ_1, χ_2. One would then hope to be able to replace the original expression $r(L_1(\mathbf{x})L_2(\mathbf{x}))r(L_3(\mathbf{x})L_4(\mathbf{x}))$ by $r_1(L_1(\mathbf{x})L_2(\mathbf{x}))r_2(L_3(\mathbf{x})L_4(\mathbf{x}))$, where

$$r_i(m) = 4\sum_{d|m} \chi_i(m) \quad (i = 1, 2).$$

If one also imposed congruence restrictions on the values of the forms $L_j(\mathbf{x})$, one would then be able to count the representations of $L_1(\mathbf{x})L_2(\mathbf{x})$ and $L_3(\mathbf{x})L_4(\mathbf{x})$ by individual genera of quadratic forms. However, while these generalizations look plausible, we have checked none of the details, and make no claim as to the results one might obtain.

2 The level of distribution

In this section we shall investigate the distribution of points $\mathbf{x}$ in subsets of $\mathcal{R}_4$, subject to a set of simultaneous divisibility conditions $d_i \mid L_i(\mathbf{x})$ for $1 \le i \le 4$. Naturally, we shall only be interested in odd values of d_i. If we write $\mathbf{d} = (d_1, d_2, d_3, d_4)$, it is clear that

$$\left\{\mathbf{x} \in \mathbb{Z}^2 : d_i \mid L_i(\mathbf{x}),\ 1 \le i \le 4\right\} = \Lambda_{\mathbf{d}},$$

say, is a lattice in $\mathbb{Z}^2$. We set

$$\rho(\mathbf{d}) = \det(\Lambda_{\mathbf{d}})$$

as in the statement of Theorem 1. We note that

$$\rho(\mathbf{d}) = [\mathbb{Z}^2 : \Lambda_{\mathbf{d}}] \mid d_1 d_2 d_3 d_4. \tag{2.1}$$

We shall consider convex regions $\mathcal{R}(\mathbf{d}) \subseteq \mathcal{R}$ for which $\mathcal{R}(\mathbf{d})$ is also the interior of a simple, piecewise continuously differentiable closed curve. We will write $\partial\mathcal{R}(\mathbf{d})$ for the length of the boundary curve defining $\mathcal{R}(\mathbf{d})$ and we set

$$\mathcal{R}_4(\mathbf{d}) = \left\{\mathbf{x} \in \mathcal{R}(\mathbf{d}) : x_1 \equiv 1 \pmod 4\right\}.$$

Since $\mathcal{R}(\mathbf{d}) \subseteq \mathcal{R} \subseteq [-cX, cX]^2$ for some constant c, by part (ii) of **NC1**, we deduce that

$$\partial\mathcal{R}(\mathbf{d}) \le 8cX,$$

since $\mathcal{R}(\mathbf{d})$ is convex. We may now state our basic result on the level of distribution of a set of linear forms L_i.

Lemma 2.1 *Let $Q_1, Q_2, Q_3, Q_4 \geq 2$, and write*

$$Q = \max Q_i \quad \textit{and} \quad V = Q_1 Q_2 Q_3 Q_4.$$

Then there is an absolute constant A such that

$$\sum_{d_i \leq Q_i} \left| \#\big(\Lambda_{\mathbf{d}} \cap \mathcal{R}_4(\mathbf{d})\big) - \frac{\operatorname{meas}(\mathcal{R}(\mathbf{d}))}{4\rho(d)} \right| \ll (XV^{1/2} + XQ + V)(\log Q)^A,$$

where the d_i run over odd integers.

A very similar result is proved by Daniel [4, Lemma 3.2], to which we refer the reader for details. As in [4, (3.11)] we find that

$$\left| \#(\Lambda_{\mathbf{d}} \cap \mathcal{R}(\mathbf{d})) - \frac{\operatorname{meas}(\mathcal{R}(\mathbf{d}))}{\rho(d)} \right| \ll \frac{\partial \mathcal{R}(\mathbf{d})}{|\mathbf{v}|} + 1 \ll \frac{X}{|\mathbf{v}|} + 1,$$

for some nonzero vector $\mathbf{v} \in \Lambda_{\mathbf{d}}$ with coprime coordinates, satisfying

$$|\mathbf{v}| \ll \det(\Lambda_{\mathbf{d}})^{1/2}.$$

By (2.1) we then deduce that $|\mathbf{v}| \ll V^{1/2}$. A trivial modification of Daniel's argument yields

$$\left| \#(\Lambda_{\mathbf{d}} \cap \mathcal{R}_4(\mathbf{d})) - \frac{\operatorname{meas}(\mathcal{R}(\mathbf{d}))}{4\rho(d)} \right| \ll \frac{X}{|\mathbf{v}|} + 1.$$

When none of the forms $L_i(\mathbf{v})$ vanish, we may estimate

$$\sum_{d_1, d_2, d_3, d_4 \leq Q} |\mathbf{v}|^{-1} \tag{2.2}$$

exactly as in [4, §3], giving a bound $O(V^{1/2}(\log Q)^A)$. However if $L_i(\mathbf{v}) = 0$ for some i we must argue differently. (This situation does not arise in Daniel's work since he has an irreducible form f of degree $k > 1$, so that $f(\mathbf{v})$ cannot vanish.) Since $\mathbf{v}$ has coprime coordinates, there can be only two possibilities for $\mathbf{v}$ for each value of i. Thus we will have $|\mathbf{v}| \ll 1$, with a constant depending only on the forms L_i. Moreover, if $L_i(\mathbf{v}) = 0$ we then have $0 \neq L_j(\mathbf{v}) \ll 1$ for $j \neq i$. Thus d_i may take any value up to Q, while for $j \neq i$ there are only $O(1)$ available values for d_j. It follows that vectors $\mathbf{v}$ for which some $L_i(\mathbf{v})$ vanishes will contribute $O(Q_i)$ to (2.2). This is sufficient for Lemma 2.1.

3 The leading term

In this section we shall examine the dominant contribution to the sum S given by (1.3). We shall use the fact that

$$r(n) = 4\sum_{d|n}\chi(d)$$

for any positive integer n, where

$$\chi(d) = \begin{cases} +1 & \text{if } d \equiv 1 \pmod 4, \\ -1 & \text{if } d \equiv 3 \pmod 4, \\ 0 & \text{if } d \equiv 0 \pmod 2). \end{cases}$$

Since $L_i(\mathbf{x}) > 0$ and $L_i(\mathbf{x}) \equiv 1 \pmod 4$ in our situation, we have

$$\begin{aligned} r(L_i(\mathbf{x})) = 4\sum_{d|L_i(\mathbf{x})}\chi(d) &= 4\sum_{\substack{d|L_i(\mathbf{x})\\ d\le X^{1/2}}}\chi(d) + 4\sum_{\substack{d|L_i(\mathbf{x})\\ d>X^{1/2}}}\chi(d) \\ &= 4\sum_{\substack{d|L_i(\mathbf{x})\\ d\le X^{1/2}}}\chi(d) + 4\sum_{\substack{L_i(\mathbf{x})=ed\\ d>X^{1/2}}}\chi(d) \\ &= 4\sum_{\substack{d|L_i(\mathbf{x})\\ d\le X^{1/2}}}\chi(d) + 4\sum_{\substack{L_i(\mathbf{x})=ed\\ d>X^{1/2}}}\chi(e) \\ &= 4\sum_{\substack{d|L_i(\mathbf{x})\\ d\le X^{1/2}}}\chi(d) + 4\sum_{\substack{e|L_i(\mathbf{x})\\ L_i(\mathbf{x})>eX^{1/2}}}\chi(e) \\ &= 4A_+(L_i(\mathbf{x})) + 4A_-(L_i(\mathbf{x})), \end{aligned} \tag{3.1}$$

say. We shall use this decomposition for the terms corresponding to L_1, L_2, L_3, and for L_4 we shall write similarly

$$r(L_4(\mathbf{x})) = 4B_+(L_4(\mathbf{x})) + 4C(L_4(\mathbf{x})) + 4B_-(L_4(\mathbf{x})),$$

where

$$\begin{gathered} B_+(m) = \sum_{\substack{d|m\\ d\le Y}}\chi(d), \quad C(m) = \sum_{\substack{d|m\\ Y<d\le X/Y}}\chi(d), \\ \text{and} \quad B_-(m) = \sum_{\substack{e|m\\ m>eX/Y}}\chi(e). \end{gathered} \tag{3.2}$$

Here $Y \leq X^{1/2}$ is a parameter to be specified in due course. For the sums A_- and B_- we note that if $\mathbf{x}$ is confined to a region $\mathcal{R}$ satisfying part (iii) of **NC1**, then the variables e which occur in the defining sums will satisfy $e \ll X^{1/2}$ and $e \ll Y$ in the two cases respectively.

We now write

$$S = \sum_{\mathbf{x} \in \mathcal{R}_4} r(L_1(\mathbf{x}))r(L_2(\mathbf{x}))r(L_3(\mathbf{x}))r(L_4(\mathbf{x}))$$

in the form

$$4S_+ + 4S_- + 4S_0,$$

where

$$S_\pm = \sum_{\mathbf{x} \in \mathcal{R}_4} r(L_1(\mathbf{x}))r(L_2(\mathbf{x}))r(L_3(\mathbf{x}))B_\pm(L_4(\mathbf{x}))$$

$$\text{and} \quad S_0 = \sum_{\mathbf{x} \in \mathcal{R}_4} r(L_1(\mathbf{x}))r(L_2(\mathbf{x}))r(L_3(\mathbf{x}))C(L_4(\mathbf{x})). \tag{3.3}$$

For the sums $S_\pm$ we shall use the decomposition (3.1) to produce a total of 8 subsums

$$S_{\pm,\pm,\pm,\pm} = \sum_{\mathbf{x} \in \mathcal{R}_4} A_\pm(L_1(\mathbf{x}))A_\pm(L_2(\mathbf{x}))A_\pm(L_3(\mathbf{x}))B_\pm(L_4(\mathbf{x})),$$

so that

$$S = 4S_0 + 4^4 \sum S_{\pm,\pm,\pm,\pm}. \tag{3.4}$$

We shall see later that S_0 is negligible. In this section we consider the remaining terms. Each of the sums $S_{\pm,\pm,\pm,\pm}$ is treated in the same way, so we shall consider the case of $S_{+,+,-,-}$, which is typical. We shall write $Q_1 = Q_2 = X^{1/2}$, and take

$$Q_3 = c_3 X^{1/2} \quad \text{and} \quad Q_4 = c_4 Y,$$

with suitable constants c_3 and c_4, so that the variables e in the sums for $A_-(L_3(\mathbf{x}))$ and $B_-(L_4(\mathbf{x}))$ will satisfy $e \leq Q_3$ and $e \leq Q_4$ respectively. With this convention, the definitions of $A_\pm$ and $B_\pm$ show that

$$S_{+,+,-,-} = \sum_{d_i \leq Q_i} \chi(d_1 d_2 d_3 d_4) \#\big(\Lambda_{\mathbf{d}} \cap \mathcal{R}_4(\mathbf{d})\big),$$

where

$$\mathcal{R}(\mathbf{d}) = \{\mathbf{x} \in \mathcal{R} : L_3(\mathbf{x}) > d_3 X^{1/2},\ L_4(\mathbf{x}) > d_4 X/Y\}. \tag{3.5}$$

Since these sets are convex, we conclude from Lemma 2.1 that

$$\begin{aligned} S_{+,+,-,-} &= \frac{1}{4}\sum_{d_i \le Q_i} \chi(d_1d_2d_3d_4)\rho^{-1}(\mathbf{d})\,\mathrm{meas}(\mathcal{R}(\mathbf{d})) \\ &\qquad + O(\{X^{7/4}Y^{1/2} + X^{3/2} + X^{3/2}Y\}(\log X)^A). \end{aligned}$$

Since $Y \le X^{1/2}$, the error term is $O(X^{7/4}Y^{1/2}(\log X)^A)$, which will be acceptable if we take

$$Y = X^{1/2}(\log X)^{-2A-2}, \tag{3.6}$$

as we now do. Thus for the general sum we have

$$S_{\pm,\pm,\pm,\pm} = \frac{1}{4}\sum_{d_i \le Q_i} \chi(d_1d_2d_3d_4)\rho^{-1}(\mathbf{d})\,\mathrm{meas}(\mathcal{R}(\mathbf{d})) + O(X^2(\log X)^{-1}). \tag{3.7}$$

We now consider the sum

$$\sum_{A_i < d_i \le B_i} \chi(d_1d_2d_3d_4)\rho^{-1}(\mathbf{d}), \tag{3.8}$$

where $B_i \le 2A_i$ for $1 \le i \le 4$. We may suppose without loss of generality that

$$A_4 \ge A_1, A_2, A_3. \tag{3.9}$$

We shall require some information on the function $\rho(\mathbf{d})$. By the Chinese Remainder Theorem there is a multiplicative property

$$\rho(d_1e_1, \ldots, d_4e_4) = \rho(d_1, \ldots, d_4)\rho(e_1, \ldots, e_4), \tag{3.10}$$

whenever

$$\mathrm{hcf}(d_1d_2d_3d_4, e_1e_2e_3e_4) = 1.$$

For most primes it is easy to handle the function ρ explicitly. As in the introduction, we write Δ for the product of the 6 possible 2×2 determinants Δ_{ij} formed from the various pairs L_i, L_j of forms. Thus if p is a prime which does not divide Δ, then for any pair $i \ne j$, we see that $p \mid L_i(\mathbf{x}), L_j(\mathbf{x})$ implies $p \mid \mathbf{x}$. Hence if

$$p^{e_i} \mid L_i(\mathbf{x}) \quad (1 \le i \le 4) \tag{3.11}$$

for a prime $p \nmid \Delta$, and $e_{\sigma(1)} \ge e_{\sigma(2)} \ge e_{\sigma(3)} \ge e_{\sigma(4)}$ for some permutation σ, then (3.11) is equivalent to

$$p^{e_{\sigma(2)}} \mid \mathbf{x} \quad \text{and} \quad p^{e_{\sigma(1)} - e_{\sigma(2)}} \mid L_{\sigma(1)}(p^{-e_{\sigma(2)}}\mathbf{x}).$$

Thus

$$\rho(p^{e_1},\dots,p^{e_4}) = p^{e_{\sigma(1)}+e_{\sigma(2)}}, \quad p\nmid\Delta. \tag{3.12}$$

For primes $p\mid\Delta$ we conclude similarly that

$$\rho(p^{e_1},\dots,p^{e_4}) \gg_{\Delta} p^{e_{\sigma(1)}+e_{\sigma(2)}}. \tag{3.13}$$

Turning to (3.8) we set $f = d_1d_2d_3\Delta$, and we write $d_4 = gh$, where

$$g = \prod_{p^e\,\|\,d_4,\,p\mid f} p^e, \quad \text{and} \quad (h,f)=1.$$

Then

$$\sum_{A_4<d_4\le B_4} \chi(d_4)\rho^{-1}(\mathbf{d}) = \sum_{g\le B_4}\chi(g)\rho^{-1}(d_1,d_2,d_3,g)\sum_{\substack{A_4/g<h\le B_4/g\\(h,f)=1}}\chi(h)\rho^{-1}(1,1,1,h).$$

In view of (3.12) we see that the inner sum is

$$\begin{aligned}\sum_{\substack{A_4/g<h\le B_4/g\\(h,f)=1}}\chi(h)/h &= \sum_{d\mid f}\mu(d)\sum_{\substack{A_4/g<h\le B_4/g\\ d\mid h}}\chi(h)/h\\ &= \sum_{d\mid f}\mu(d)\chi(d)/d\sum_{A_4/gd<j\le B_4/gd}\chi(j)/j.\end{aligned}$$

However

$$\sum_{J<j\le K}\chi(j)/j \ll J^{-1},$$

so the sum above is $O(gf^{\varepsilon}A_4^{-1})$, for any $\varepsilon>0$.

It follows that (3.8) is

$$\ll A_4^{-1}\sum_{d_1,d_2,d_3}\sum_{g\le B_4}(d_1d_2d_3)^{\varepsilon}g\rho^{-1}(d_1,d_2,d_3,g). \tag{3.14}$$

We shall estimate this sum by Rankin's method. For any fixed $\delta>0$ we have

$$d_i^{\varepsilon} \ll d_i^{\delta} \ll A_i^{2\delta} d_i^{-\delta}$$

providing that ε is small enough. Similarly we have

$$1 \ll A_4^{\delta} g^{-\delta}.$$

It follows that

$$\begin{aligned}
&\sum_{d_1,d_2,d_3} \sum_{g \le B_4} g(d_1d_2d_3)^{\varepsilon} \rho(d_1,d_2,d_3,g)^{-1} \\
&\qquad \ll (A_1A_2A_3A_4)^{2\delta} \sum_{d_1,d_2,d_3} \sum_{g \le B_4} g^{1-\delta}(d_1d_2d_3)^{-\delta}\rho(d_1,d_2,d_3,g)^{-1} \\
&\qquad \ll (A_1A_2A_3A_4)^{2\delta} \sum_{d_1,d_2,d_3=1}^{\infty} \sum_{g=1}^{\infty} g^{1-\delta}(d_1d_2d_3)^{-\delta}\rho(d_1,d_2,d_3,g)^{-1}, \qquad (3.15)
\end{aligned}$$

where g is still restricted to integers composed solely of prime factors p dividing $f = \Delta d_1d_2d_3$. In view of the multiplicative property (3.10) we can factorize the 4-fold sum on the right. For each prime p we write $d_1 = p^a$, $d_2 = p^b$, $d_3 = p^c$ and $g = p^d$, so that the corresponding factor is

$$\sum_{a,b,c,d=0}^{\infty} p^{d-(a+b+c+d)\delta}\rho(p^a,p^b,p^c,p^d)^{-1}, \qquad (3.16)$$

subject to the condition that if $p \nmid \Delta$ then there are no terms with $a = b = c = 0$ and $d > 0$. For those primes p which do not divide Δ the above sum is $1 + O(\Sigma_p)$, where Σ_p is a sum of the form

$$\begin{aligned}
\sum_{a=1}^{\infty} \sum_{0 \le b,c \le a} \sum_{d=0}^{\infty} & p^{d-(a+b+c+d)\delta}\rho(p^a,p^b,p^c,p^d)^{-1} \\
&\le \sum_{a=1}^{\infty} \sum_{0 \le b,c \le a} \sum_{d=0}^{\infty} p^{d-(a+b+c+d)\delta} p^{-a-d} \\
&\le p^{-1-\delta} \left\{ \sum_{e=0}^{\infty} p^{-e\delta} \right\}^4 = O_{\delta}(p^{-1-\delta}),
\end{aligned}$$

by (3.12). The product of all such factors (3.16) is therefore $O_{\delta}(1)$. For the remaining primes we use (3.13) to show similarly that (3.16) is $O_{\delta,\Delta}(1)$. The 4-fold sum in (3.16) is therefore bounded, and on choosing $\delta = 1/10$, say, we see from (3.9) that (3.15) is $O(A_4^{4/5})$, and hence, from (3.14) that

$$\sum_{A_i < d_i \le B_i} \chi(d_1d_2d_3d_4)\rho^{-1}(\mathbf{d}) \ll (A_1A_2A_3A_4)^{-1/20}.$$

We may now use repeated summation by parts to show that

$$\sum_{d_i \le A_i} \chi(d_1d_2d_3d_4)\rho^{-1}(\mathbf{d})(d_1d_2d_3d_4)^{-\delta} = S(\delta) + O((\min A_i)^{-1/20}) \qquad (3.17)$$

uniformly for $\delta > 0$, with

$$S(\delta) = \sum_{d_1,d_2,d_3,d_4=1}^{\infty} \chi(d_1d_2d_3d_4)\rho^{-1}(\mathbf{d})(d_1d_2d_3d_4)^{-\delta}.$$

The sum $S(\delta)$ is absolutely convergent for such δ. Indeed by (3.10) it suffices to consider the behaviour of the various Euler factors. For each prime the corresponding factor is

$$\sum_{a,b,c,d=0}^{\infty} \chi(p)^{a+b+c+d} p^{-(a+b+c+d)\delta} \rho(p^a, p^b, p^c, p^d)^{-1} = E_p(\delta), \qquad (3.18)$$

say. We write this in the form $1 + \Sigma$ where

$$\Sigma \ll \sum_{a=1}^{\infty} \sum_{b,c,d=0}^{\infty} p^{-a-(a+b+c+d)\delta} \ll p^{-1-\delta},$$

by (3.12) and (3.13). This suffices to ensure absolute convergence for $\delta > 0$. Similarly, when $p \nmid \Delta$ we have $\rho(p, 1, 1, 1) = p$ by (3.12), whence

$$E_p(\delta) = 1 + 4\chi(p)/p^{-1-\delta} + O(p^{-2}) = \{1 - \chi(p)/p^{1+\delta}\}^{-4}\{1 + O(p^{-2})\},$$

uniformly for $\delta > 0$. It follows that we can write $S(\delta) = L(1 + \delta, \chi)^4 F(1 + \delta)$ where

$$F(s) = \prod_p E_p(s-1)\{1 - \chi(p)p^{-s}\}^4 \qquad (3.19)$$

is absolutely and uniformly convergent for $\mathrm{Re}(s) \ge 1$. This allows us to take the limit in (3.17) as δ tends to zero, so that

$$\sum_{d_i \le A_i} \chi(d_1d_2d_3d_4)\rho^{-1}(\mathbf{d}) = \left(\frac{\pi}{4}\right)^4 F(1) + O((\min A_i)^{-1/20}).$$

It remains to introduce the factor $\mathrm{meas}(\mathcal{R}(\mathbf{d}))$ into this sum, which we proceed to do via partial summation. Recall that we are working with the example (3.5). For ease of notation we shall set $d_3 = x, d_4 = y$ and $f(x, y) =$

meas$(\mathcal{R}(\mathbf{d}))$. Then

$$\sum_{d_i\le Q_i}\chi(d_1d_2d_3d_4)\rho^{-1}(\mathbf{d})\,\mathrm{meas}(\mathcal{R}(\mathbf{d}))=$$
$$\int_0^{Q_3}\int_0^{Q_4} f_{xy}(x,y)\sum_{\substack{d_1\le Q_1,\, d_2\le Q_2\\ d_3\le x,\, d_4\le y}}\chi(d_1d_2d_3d_4)\rho^{-1}(\mathbf{d})dxdy$$

by partial summation, on noting that $f(Q_3,y)=f(x,Q_4)=0$ for all x,y. We therefore obtain

$$S_{+,+,-,-}=\frac{1}{4}\left(\frac{\pi}{4}\right)^4 F(1)\,\mathrm{meas}\,\mathcal{R}+O\Big(\int_0^{Q_3}\int_0^{Q_4}|f_{xy}(x,y)|(\min(x,y))^{-1/20}\Big).$$

However $F_{xy}(x,y)\ll X^2/Q_3Q_4$, as one sees from (3.15). Hence the error term above is $O(X^2(\min Q_i)^{-1/20})$. We therefore deduce that

$$S_{+,+,-,-}=\frac{1}{4}\left(\frac{\pi}{4}\right)^4 F(1)\,\mathrm{meas}\,\mathcal{R}+O(X^{79/40}(\log X)^A),$$

and similarly for each of the sums $S_{\pm,\pm,\pm,\pm}$. If we now refer to (3.4) and (3.7), we may conclude as follows.

Lemma 3.1 *We have*

$$S=4\pi^4F(1)\,\mathrm{meas}\,\mathcal{R}+4S_0+O(X^2(\log X)^{-1}),$$

where $F(1)$ is given by (3.18) and (3.19), and S_0 is given by (3.3).

4 The sum S_0—first steps

Clearly we have

$$\begin{aligned} S_0 &\ll \sum_{\mathbf{x}\in\mathcal{R}_4} r(L_1(\mathbf{x}))r(L_2(\mathbf{x}))r(L_3(\mathbf{x}))|C(L_4(\mathbf{x}))| \\ &= \sum_{m\in\mathcal{B}} S_0(m)|C(m)|, \end{aligned} \tag{4.1}$$

where

$$\mathcal{B}=\left\{m\in\mathbb{Z}:\begin{matrix}\exists d\mid m \text{ s.t.}\\ Y<d\le X/Y\end{matrix}\right\}\cap\left\{m\in\mathbb{Z}:\begin{matrix}\exists \mathbf{x}\in\mathcal{R}_4 \text{ s.t.}\\ L_4(\mathbf{x})=m\end{matrix}\right\} \tag{4.2}$$

and

$$S_0(m) = \sum_{\mathbf{x} \in \mathcal{A}(m)} r(L_1(\mathbf{x}))r(L_2(\mathbf{x}))r(L_3(\mathbf{x}))$$

with

$$\mathcal{A}(m) = \{\mathbf{x} \in \mathcal{R}_4 : L_4(\mathbf{x}) = m\}.$$

Suppose that the forms L_i are given by

$$L_i(x_1, x_2) = A_i x_1 + B_i x_2, \quad (1 \leq i \leq 4). \tag{4.3}$$

We have arranged that $L_i(x_1, x_2) \equiv 1 \pmod 4$ whenever we have $x_1, x_2 \in \mathbb{Z}$ and $x_1 \equiv 1 \pmod 4$. It follows that $A_i \equiv 1 \pmod 4$ and $B_i \equiv 0 \pmod 4$. In particular $A_i \neq 0$. If we now substitute $m = L_4(\mathbf{x})$ for x_1, so that $x_1 = (m - B_4 x_2)/A_4$, and write $x_2 = n$ for ease of notation, we find that

$$L_i(\mathbf{x}) = \frac{a_i m + b_i n}{A_4} = L_i'(m, n),$$

say, where

$$a_i = A_i, \quad b_i = A_4 B_i - B_4 A_i, \quad (1 \leq i \leq 3).$$

Thus we have

$$a_i \equiv 1 \pmod 4, \quad b_i \equiv 0 \pmod 4, \quad (1 \leq i \leq 3). \tag{4.4}$$

Note that, as $\mathbf{x}$ runs over $\mathbb{Z}^2$, not every value $m \in \mathbb{Z}$ need occur. Indeed, since $x_1 \equiv 1 \pmod 4$ we will have $m \equiv 1 \pmod 4$. We also observe that if $\mathbf{x}$ runs over $\mathcal{R}$, then the corresponding values of m and n will satisfy $m, n \ll X$. Finally we note that we can clear the denominator in L_i', so that $r(L_i'(m,n)) \leq r(A_4(a_i m + b_i n))$.

We now write

$$H = 6\Delta A_4^3 \prod_{1 \leq i \leq 3} b_i.$$

This will be nonzero since no two of the original forms $L_1, \ldots, L_4$ were proportional. We also define a multiplicative function $r_1(n)$ by setting

$$r_1(p^e) = \begin{cases} (e+1)^3, & \text{if } p \mid H \text{ or } e \geq 2, \\ 1 + \chi(p), & \text{otherwise.} \end{cases}$$

Using the multiplicative property of the function $r(n)$ one can then verify that

$$r(L_1'(m,n))r(L_2'(m,n))r(L_3'(m,n)) \leq 64 r_1(F(n)),$$

where

$$F(n) = A_4^3 \prod_{i=1}^{3} (a_i m + b_i n). \tag{4.5}$$

Our principal tool in handling $S_0(m)$ will be a theorem of Nair [7], which will provide an upper bound of the correct order of magnitude. In order to apply Nair's result we must remove fixed prime factors from F. Thus we first write $F(X) = cG(X)$, where $G(X)$ is a primitive integer polynomial, and $c \mid H$. It follows that $r_1(F(n)) \ll r_1(G(n))$. The only fixed prime factors that a primitive cubic polynomial can have are $p = 2$ and $p = 3$. However since $m \equiv 1 \pmod 4$ we see from (4.5) that $p = 2$ can never divide $G(n)$. If $G(X)$ has $p = 3$ as a fixed prime divisor then $G(X) \equiv \pm(X^3 - X) \pmod 3$. Thus if we split the integers n into the three possible congruence classes $n \equiv n_0 \pmod 3$, and write $n = 3\widehat{n} + n_0$ we see that

$$\frac{G(n)}{3} = 9\frac{G'''(n_0)}{6}\widehat{n}^3 + 3\frac{G''(n_0)}{2}\widehat{n}^2 + G'(n_0)\widehat{n} + \frac{1}{3}G(n_0) = \widehat{G}(\widehat{n}),$$

say. Since $G'(n_0) \equiv \mp 1 \pmod 3$ we see that $\widehat{G}$ does not have $p = 3$ as a fixed prime divisor. Thus, by splitting the range for n into three congruence classes if necessary, we can produce a polynomial with no fixed prime divisor.

We now state the following special case of Nair's theorem [7].

Lemma 4.1 *Let $f(n)$ be a nonnegative multiplicative function satisfying the bound $f(p^e) \le (e+1)^4$ for every prime power p^e. Let $G(X) \in \mathbb{Z}[X]$ be a polynomial of degree at most 4, without repeated roots, and with no fixed prime factor. Write $\rho(p)$ for the number of roots of G modulo p, and $||G||$ for the sum of the moduli of the coefficients of G. Then for any $\delta > 0$ there is a constant c_δ such that*

$$\sum_{n \le N,\, G(n) > 0} f(G(n)) \ll_\delta N \prod_{p \le N}\Big(1 - \frac{\rho(p)}{p}\Big) \exp\left(\sum_{p \le N} \frac{f(p)\rho(p)}{p}\right),$$

for $N \ge c_\delta ||G||^\delta$.

For our application the range for n will be an interval of length $N \ll X$, which will have to be translated by a distance $O(X)$ in order to produce the interval $(0, N]$. This has the effect of modifying the coefficients of the original polynomial G. However even after this translation we will have $||G|| \ll X^3$. Given the form (4.5) of F we see that G will have three linear factors. Moreover we have $\rho(p) = 1$ for $p \mid m$, while if $p \nmid mH$ we will have $\rho(p) = 3$, since $p \mid a_i b_j - a_j b_i$ would imply $p \mid \Delta$. We will therefore have

$$\begin{aligned}
S_0(m) &\ll \sum_{n \le N,\, G(n)>0} r_1(G(n)) \\
&\ll N \prod_{p \le N} \Big(1 - \frac{\rho(p)}{p}\Big) \exp\left(\sum_{p \le N} \frac{r_1(p)\rho(p)}{p}\right) \\
&\ll N \prod_{3<p \le N} \Big(1 - \frac{\rho(p)}{p}\Big) \exp\left(\sum_{p \le N} \frac{3r_1(p)}{p}\right) \\
&\ll N \prod_{p \mid m,\, p>3} \frac{1-1/p}{1-3/p} \prod_{3<p \le N} \Big(1 - \frac{3}{p}\Big) \exp\left(\sum_{p \le N} \frac{3r_1(p)}{p}\right) \\
&\ll N \left(\frac{\sigma(m)}{m}\right)^2 \\
&\ll N(\log\log N)^2, \qquad (4.6)
\end{aligned}$$

providing that $N \gg_\delta X^{3\delta}$. (Here $\sigma(m)$ is the usual sum of divisors function.) Since we trivially have $r_1(G(n)) \ll X^{1/2}$ we see on taking $\delta = 1/6$ that $S_0(m) \ll X(\log\log X)^2$ whether $N \gg X^{1/2}$ or not. We therefore deduce the following result from (4.1).

Lemma 4.2 *We have*

$$S_0 \ll X(\log\log X)^2 \sum_{m \in \mathcal{B}} |C(m)|,$$

where $\mathcal{B}$ and $C(m)$ are given by (4.2) and (3.2) respectively.

5 Completion of the proof of Theorem 1

Cauchy's inequality shows that

$$\sum_{m \in \mathcal{B}} |C(m)| \le (\#\mathcal{B})^{1/2} \Big(\sum_{1 \le m \ll X} |C(m)|^2 \Big)^{1/2}. \qquad (5.1)$$

However it is clear that if we let M and D run over powers of 2, then

$$\begin{aligned}
\#\mathcal{B} &\le \#\big\{m \ll X : \exists d \mid m,\ Y < d \le X/Y\big\} \\
&\ll \log(X/Y^2) \sum_M \#\big\{M < m \le 2M : \exists d \mid m,\ D < d \le 2D\big\} \qquad (5.2)
\end{aligned}$$

for some D in the range $Y \ll D \ll X/Y$. Clearly we may replace d by m/d, so that

$$\#\big\{M < m \le 2M : \exists d \mid m,\ D < d \le 2D\big\} \\ \le \#\big\{M < m \le 2M : \exists d \mid m,\ M/2D < d < 2M/D\big\}.$$

Now we may apply the following result.

Lemma 5.1 *We have*

$$\#\big\{n \le x : \exists d \mid n,\ y < d \le 2y\big\} \ll \frac{x}{(\log y)^{\eta}(\log\log y)^{1/2}}$$

uniformly for $3 \le y \le x$, *where* η *is given by (1.5).*

This is the case $u = 1$, $\beta = 0$ of Theorem 21, part (ii) in Hall and Tenenbaum, see [5, (2.2) and (2.3)].

Lemma 5.1 yields

$$\#\{M < m \le 2M : \exists d \mid m,\ D < d \le 2D\} \ll \frac{M}{(\log X)^{\eta}(\log\log X)^{1/2}}$$

whenever $M \ge X^{3/4}$. For smaller values of M we merely use the trivial bound $O(M)$. Then (5.2) and (3.6) imply that

$$\#\mathcal{B} \ll X(\log X)^{-\eta}(\log\log X)^{1/2}. \tag{5.3}$$

It remains to consider

$$\sum_{1 \le m \le cX} |C(m)|^2,$$

for a suitable constant c. We expand the term $|C(m)|^2$ and write $(d_1, d_2) = h$ and $d_i = hk_i$ to produce

$$\begin{aligned}\sum_{1 \le m \le cX} |C(m)|^2 \\ &= \sum_{d_1,d_2 \in (Y,X/Y]} \chi(d_1 d_2)\#\big\{m \le cX : [d_1, d_2] \mid m\big\} \\ &= \sum_{h \le X/Y}\ \sum_{\substack{k_1,k_2 \in (Y/h, X/Yh] \\ (k_1,k_2)=1}} \chi(h^2 k_1 k_2)\#\big\{m \le cX : hk_1k_2 \mid m\big\} \\ &= \sum_{h \le X/Y}\ \sum_{\substack{k_1,k_2 \in (Y/h, X/Yh] \\ (k_1,k_2)=1}} \chi(h^2 k_1 k_2)\#\big\{n \le cX/hk_1k_2\big\} \\ &= \sum_{h \le X/Y}\ \sum_{k_1 \in (Y/h, X/Yh]} \chi(h^2 k_1) \sum_{n \le \min(cX/Yk_1, cX/hk_1)}\ \sum_{k_2} \chi(k_2), \end{aligned} \tag{5.4}$$

where the innermost sum in the final expression is subject to the conditions $Y/h < k_2 \le \min(X/Yh, cX/hk_1n)$ and $(k_2, k_1) = 1$.

In general we have

$$\begin{aligned}
\sum_{k\le K,\,(k,s)=1} \chi(k) &= \sum_{d|s}\mu(d) \sum_{k\le K,\, d|k} \chi(k)\\
&= \sum_{d|s}\mu(d)\chi(d) \sum_{j\le K/d} \chi(j)\\
&\ll \sum_{d|s} |\mu(d)\chi(d)|\\
&\ll \tau(s),
\end{aligned}$$

where τ is the usual divisor function. Inserting this bound into (5.4) we deduce that

$$\begin{aligned}
\sum_{1\le m\le cX} |C(m)|^2 &\ll \sum_{h\le X/Y}\ \sum_{k_1\in(Y/h,X/Yh]}\ \sum_{n\le\min(cX/Yk_1,cX/hk_1)} \tau(k_1)\\
&\ll \sum_{h\le X/Y}\ \sum_{k_1\in(Y/h,X/Yh]} \min\Big(\frac{X}{Yk_1}, \frac{X}{hk_1}\Big)\tau(k_1)\\
&= \sum_{h\le X/Y} \min\Big(\frac{X}{Y}, \frac{X}{h}\Big) \sum_{k_1\in(Y/h,X/Yh]} \tau(k_1)/k_1\\
&\ll \sum_{h\le X/Y} \min\Big(\frac{X}{Y}, \frac{X}{h}\Big) \log^2(X/Yh)\\
&\ll XY^{-1}\sum_{h\le Y} \log^2(X/Yh) + X \sum_{Y<h\le X/Y} h^{-1}\log^2(X/Yh)\\
&\ll X\log^2(XY^{-2}) + X\log^3(XY^{-2}).
\end{aligned}$$

Our choice (3.6) of Y then ensures that

$$\sum_{1\le m\le cX} |C(m)|^2 \ll X(\log\log X)^3,$$

so that (5.1), (5.2) and Lemma 4.2 produce the bound

$$S_0 \ll X^2(\log X)^{-\eta/2}(\log\log X)^{15/4}.$$

This suffices, in conjunction with Lemma 3.1, for Theorem 1.

6 Proof of Theorem 2—preliminaries

Our starting point for the proof of Theorem 2 is the identity

$$r(mn) = \frac{1}{4}\sum_{d|m,n} \mu(d)\chi(d)r(m/d)r(n/d),$$

valid for any positive integers m, n. This identity allows us to pass from a problem about solutions of a single equation $mn = r^2 + s^2$ to one which involves a series of systems $m = d(t^2 + u^2)$, $n = d(v^2 + w^2)$ for varying d. One can think of this as corresponding to a simple 'descent' process.

In view of part (iii) of **NC2**, we may take $m = L_1, n = L_2$, or alternatively $m = L_3, n = L_4$ in the above identity. Thus, if

$$S = \sum_{\mathbf{x}\in\mathcal{R}_2} r(L_1(\mathbf{x})L_2(\mathbf{x}))r(L_3(\mathbf{x})L_4(\mathbf{x})),$$

we have

$$S = \frac{1}{16}\sum_{d,d'} \mu(d)\mu(d')\chi(dd') \times$$
$$\sum_{\mathbf{x}\in\mathcal{R}_2} r(L_1(\mathbf{x})/d)r(L_2(\mathbf{x})/d)r(L_3(\mathbf{x})/d')r(L_4(\mathbf{x})/d'),$$

where we set $r(q) = 0$ if q is not an integer. Since L_i is always odd for $\mathbf{x} \in \mathcal{R}_2$, part (iv) of **NC2** shows that we must have $x_1 \equiv \nu d \pmod 4$ if $r(L_1/d) \neq 0$, and similarly $x_1 \equiv \nu' d' \pmod 4$ if $r(L_3/d) \neq 0$. In particular, only terms for which $dd' \equiv \nu\nu' \pmod 4$ make a nonzero contribution, so that

$$S = \frac{\chi(\nu\nu')}{16} \sum_{dd'\equiv\nu\nu' \pmod 4} \mu(d)\mu(d')S(d,d'), \tag{6.1}$$

where

$$S(d,d') = \sum_{\mathbf{x}\in\mathcal{R},\, x_1\equiv\nu d \pmod 4} r(L_1(\mathbf{x})/d)r(L_2(\mathbf{x})/d)r(L_3(\mathbf{x})/d')r(L_4(\mathbf{x})/d').$$

Henceforth we shall assume, as we clearly may, that d and d' are both odd.

We now show that it suffices to establish an asymptotic formula for each individual sum $S(d, d')$.

Lemma 6.1 *Suppose that*

$$S(d,d') \ll X^2\tau(d)^5\tau(d')^5[d,d']^{-2} \tag{6.2}$$

uniformly for all square-free d, d', where $[d, d']$ denotes the least common multiple of d and d'. Assume further that

$$S(d,d') = C(d,d')\,\mathrm{meas}\,\mathcal{R} + o(X^2) \tag{6.3}$$

for all fixed square-free d, d', and that

$$C(d,d') \ll \tau(d)^5\tau(d')^5[d,d']^{-2} \tag{6.4}$$

for square-free d, d'. Then, under **NC2**, *we have*

$$S = C \operatorname{meas} \mathcal{R} + o(X^2), \tag{6.5}$$

with

$$C = \frac{\chi(\nu\nu')}{16} \sum_{dd' \equiv \nu\nu' \pmod 4} \mu(d)\mu(d')C(d, d'). \tag{6.6}$$

Notice that we do not require any uniformity in d, d' for (6.3). It suffices that (6.3) should hold for each fixed pair d, d'.

To prove the lemma we set

$$E(d, d'; X) = X^{-2}|S(d, d') - C(d, d') \operatorname{meas} \mathcal{R}|,$$

so that (6.2) and (6.4) yield

$$E(d, d'; X) \ll \tau(d)^5 \tau(d')^5 [d, d']^{-2}$$

uniformly in X. On the other hand, for fixed d, d' we will have $E(d, d'; X) \to 0$ as $X \to \infty$. The required result will therefore follow from the dominated convergence of the double sum

$$\sum_{d,d'=1}^{\infty} E(d, d'; X),$$

providing that we can show that

$$\sum_{d,d'=1}^{\infty} \tau(d)^5 \tau(d')^5 [d, d']^{-2}$$

converges. However if we set $(d, d') = h$ and $d = hk, d' = hk'$ we will have

$$\sum_{d,d'=1}^{\infty} \tau(d)^5 \tau(d')^5 [d, d']^{-2} \leq \sum_{h,k,k'=1}^{\infty} \tau(k)^5 \tau(k')^5 \tau(h)^{10} (hkk')^{-2},$$

and the required result follows.

We now establish the bound (6.2), using Nair's result, Lemma 4.1. We begin by writing Δ for the product of the 6 possible 2×2 determinants formed from the various pairs L_i, L_j of forms, as previously. Thus if p is a prime which does not divide Δ, then $p \mid L_i(\mathbf{x}), L_j(\mathbf{x})$ implies $p \mid \mathbf{x}$, providing that $i \neq j$. We shall put $e = (d, \Delta), e' = (d', \Delta)$ and $f = d/e, f' = d'/e'$. If d, d' are square-free, we see that e and f are square-free and that $(f, \Delta) = 1$. Similarly e' and f' are square-free and $(f', \Delta) = 1$. The condition $d \mid L_1(\mathbf{x}), L_2(\mathbf{x})$ now implies

$f \mid \mathbf{x}$, while $d' \mid L_3(\mathbf{x}), L_4(\mathbf{x})$ implies $f' \mid \mathbf{x}$. We therefore set $\mathbf{x} = g\mathbf{y}$, where $g = [f, f']$ is the lowest common multiple of f and f'. We shall henceforth assume that $g \ll X$, as we clearly may. It now follows that

$$S(d,d') \le \sum_{\mathbf{y}} r(gL_1(\mathbf{y})/d)r(gL_2(\mathbf{y})/d)r(gL_3(\mathbf{y})/d')r(gL_4(\mathbf{y})/d'),$$

where the sum is for vectors $\mathbf{y}$ such that $g\mathbf{y} \in \mathcal{R}$ and $y_1 \equiv g\nu d \pmod 4$. If the forms L_i are given by (4.3), we conclude, using part (iv) of **NC2**, that $A_i \ne 0$ for $1 \le i \le 4$. We proceed to define a multiplicative function $r_2(n)$ by setting

$$r_2(p^e) = \begin{cases} 1+\chi(p) & \text{if } p \nmid 3dd' \prod A_i \text{ and } e = 1, \\ (1+e)^4 & \text{otherwise.} \end{cases}$$

Then

$$r(gL_1(\mathbf{y})/d)r(gL_2(\mathbf{y})/d)r(gL_3(\mathbf{y})/d')r(gL_4(\mathbf{y})/d') \\ \le 4^4\tau(g)^4 r_2\big(L_1(\mathbf{y})L_2(\mathbf{y})L_3(\mathbf{y})L_4(\mathbf{y})\big).$$

Moreover, if we regard y_2 as fixed and set $F(X) = \prod L_i(X, y_2)$, we will have $F(X) = cG(X)$ for some primitive quartic polynomial $G(X)$, with $c \mid \prod A_i$. Since we are taking the forms L_i to be fixed, it follows that

$$r(gL_1(\mathbf{y})/d)r(gL_2(\mathbf{y})/d)r(gL_3(\mathbf{y})/d')r(gL_4(\mathbf{y})/d') \ll \tau(g)^4 r_2(G(y_1)).$$

We intend to apply Lemma 4.1, and we therefore investigate possible fixed prime factors p of $H(X) = G(2X+1)$. Since G is quartic and primitive we must have $p = 2$ or $p = 3$. However, for $y_1 \equiv g\nu d \pmod 4$, we see from part (iv) of **NC2** that $F(y_1)$, and hence also $G(y_1)$, must be odd. Thus $H(0) = G(1)$ is odd. There remains the case $p = 3$. Suppose that $3 \mid H(n)$ for all $n \in \mathbb{Z}$. We split the available y into congruence classes modulo 3 and consider the three polynomials

$$H_j(X) = \frac{H(3X+j)}{3}, \quad (j = 0, 1, 2).$$

Clearly the only possible fixed prime factor of H_j is $p = 3$. We claim that if H_j does have 3 as a fixed prime factor, then H_j is divisible by 3 as a polynomial. Moreover, if we then put $H_j(X) = 3K_j(X)$ we claim that H_j does not have 3 as a fixed prime factor. To prove these assertions, suppose that there is some j such that $3 \mid H_j(n)$ for all $n \in \mathbb{Z}$. Then $9 \mid H(3n+j)$, whence $9 \mid H(j) + 3nH'(j)$ for every n. It follows that $9 \mid H(j)$ and $3 \mid H'(j)$ so that 9 divides the polynomial $H(3X+j)$. Thus $3 \mid H_j(X)$ as claimed. Moreover, if $9 \mid H_j(n)$ for every n, then $27 \mid H(3n+j)$, whence $27 \mid H(j) + 3nH'(j) + 9n^2H''(j)/2$. From this we deduce that $3 \mid H''(j)$. However we then see that

$$\begin{aligned} H(m+j) &= H(j) + mH'(j) + m^2\frac{H''(j)}{2} + m^3\frac{H^{(3)}(j)}{6} + m^4\frac{H^{(4)}(j)}{24} \\ &\equiv m^3\frac{H^{(3)}(j)}{6} + m^4\frac{H^{(4)}(j)}{24} \pmod 3. \end{aligned}$$

This produces a contradiction, since we are supposing that $H(X)$ is primitive and has 3 as a fixed prime factor.

It therefore follows that we may replace $H(X)$ if necessary by a set of 3 polynomials $H_j(X)$ or $K_j(X)$ which have no fixed prime divisor. Moreover $r_2(H(3n+j)) \le r_2(3)r_2(H_j(n))$ and $r_2(H(3n+j)) \le r_2(9)r_2(K_j(n))$, so that only a factor $O(1)$ is lost. Now, if

$$S(y_2) = \sum_y r(gL_1(y,y_2)/d)r(gL_2(y,y_2)/d)r(gL_3(y,y_2)/d')r(gL_4(y,y_2)/d'),$$

where the sum over y is subject to $g(y,y_2) \in \mathcal{R}$ and $y \equiv g\nu d \pmod 4$, we find from Lemma 4.1 that if $y_2 \ne 0$, then

$$\begin{aligned} S(y_2) &\ll \frac{X}{g}\tau(g)^4 \prod_{p\le X}\Big(1-\frac{\rho(p)}{p}\Big)\exp\Big(\sum_{p\le X}\frac{r_2(p)\rho(p)}{p}\Big) \\ &\ll \frac{X}{g}\tau(g)^4 \prod_{5<p\le N}\Big(1-\frac{4}{p}\Big)\exp\Big(\sum_{p\le X}\frac{4r_2(p)}{p}\Big)\exp\Big(\sum_{p\mid dd'y_2}\frac{64}{p}\Big) \\ &\ll \frac{X}{g}\tau(g)^4\Big(\frac{\sigma(dd')}{dd'}\Big)^{64}\Big(\frac{\sigma(|y_2|)}{|y_2|}\Big)^{64}, \end{aligned}$$

as in (4.6). We trivially have

$$S(0) \ll \sum_{y\ll X/g}\tau(y)^4 \ll X^2g^{-2}.$$

We therefore deduce that

$$\begin{aligned} S(d,d') &\ll X^2g^{-2} + Xg^{-1}\tau(g)^4\tau(dd')\sum_{1\le y_2\ll X/g}\Big(\frac{\sigma(|y_2|)}{|y_2|}\Big)^{64} \\ &\ll X^2g^{-2}\tau(g)^4\tau(dd'). \end{aligned}$$

Since $g|dd'$ and $[d,d'] \mid \Delta g$, the bound (6.2) then follows.

7 Proof of Theorem 2—the asymptotic formula

We must now establish the asymptotic formula (6.5), and analyse its main term, with a view to proving the bound (6.4). We begin by showing how Theorem 1 may be applied.

The conditions $d \mid L_1(\mathbf{x}), L_2(\mathbf{x})$ and $d' \mid L_3(\mathbf{x}), L_4(\mathbf{x})$ will hold if and only if $\mathbf{x} \in \Lambda_{(d,d,d',d')}$. We therefore take $\mathbf{a}, \mathbf{b}$ as a basis for $\Lambda_{(d,d,d',d')}$ and write $\mathbf{a} = (a_1, a_2)$ and $\mathbf{b} = (b_1, b_2)$. Since (dd', dd') is clearly in $\Lambda_{(d,d,d',d')}$, we see that at least one of a_1 and b_1 must be odd, and we can therefore take a_1 to be odd. By changing the sign of a_1 if necessary we can then assume that we have $a_1 \equiv \nu d \pmod 4$, and finally, replacing $\mathbf{b}$ by $\mathbf{b} - k\mathbf{a}$ for a suitable integer k, we can assume that $4 \mid b_1$. Having normalized the basis $\mathbf{a}, \mathbf{b}$ of $\Lambda_{(d,d,d',d')}$ in this way we set $\mathbf{x} = y_1\mathbf{a} + y_2\mathbf{b}$. Moreover we write $L_i'(\mathbf{y}) = d^{-1}L_i(y_1\mathbf{a} + y_2\mathbf{b})$ for $i = 1, 2$ and similarly $L_i'(\mathbf{y}) = d'^{-1}L_i(y_1\mathbf{a} + y_2\mathbf{b})$ for $i = 3, 4$, and we set

$$\mathcal{R}'^{(0)} = \{\mathbf{y} \in \mathbb{R}^2 : y_1\mathbf{a} + y_2\mathbf{b} \in \mathcal{R}^{(0)}\}.$$

It now follows that

$$x_1 = y_1a_1 + y_2b_1 \equiv y_1\nu d \pmod 4,$$

so that for $i = 1, 2$ the condition $L_i(\mathbf{x}) \equiv \nu x_1 \pmod 4$ becomes

$$L_i'(\mathbf{y}) \equiv d^{-1}L_i(\mathbf{x}) \equiv d^{-1}\nu x_1 \equiv y_1 \pmod 4.$$

Similarly for $i = 3, 4$ we have

$$L_i'(\mathbf{y}) \equiv d'^{-1}L_i(\mathbf{x}) \equiv d'^{-1}\nu' x_1 \equiv y_1 \pmod 4,$$

since $d\nu \equiv d'\nu' \pmod 4$ in (6.1).

It is now apparent that, for fixed d, d', the forms $L_i'(\mathbf{y})$, and the region $\mathcal{R}'^{(0)}$ satisfy **NC1**. Evidently we have $\text{meas}(\mathcal{R}') = \text{meas}\,\mathcal{R}/\rho(d,d,d',d')$. For fixed d, d' we therefore deduce that

$$S(d, d') = \frac{4\pi^4 \prod_p \sigma_p(d, d')}{\rho(d, d, d', d')}\,\text{meas}\,\mathcal{R} + o(X^2)$$

for each fixed pair d, d'. Here we have

$$\sigma_p(d, d') = E_p(d, d')\{1 - \chi(p)/p\}^4$$

where

$$E_p(d, d') = \sum_{\alpha,\beta,\gamma,\delta=0}^{\infty} \chi(p)^{\alpha+\beta+\gamma+\delta}\rho_0(p^\alpha, p^\beta, p^\gamma, p^\delta)^{-1}$$

and $\rho_0(d_1,d_2,d_3,d_4)$ is the determinant of the lattice

$$\Lambda_1=\{\mathbf{y}\in\mathbb{Z}^2: d_i\mid L_i'(\mathbf{y}),\ 1\le i\le 4\}.$$

We now observe that $\rho_0(d_1,d_2,d_3,d_4)$ is also the index of the lattice Λ_1 in $\mathbb{Z}^2$, and hence can equally be identified as the index of

$$\Lambda_2=\{\mathbf{x}=y_1\mathbf{a}+y_2\mathbf{b}:\mathbf{y}\in\mathbb{Z}^2,\ d_i\mid L_i'(\mathbf{y}),\ 1\le i\le 4\}$$

in

$$\Lambda_3=\{\mathbf{x}=y_1\mathbf{a}+y_2\mathbf{b}:\mathbf{y}\in\mathbb{Z}^2\}.$$

However we have

$$\Lambda_2=\{\mathbf{x}\in\mathbb{Z}^2: dd_1\mid L_1(\mathbf{x}),\ dd_2\mid L_2(\mathbf{x}),\ d'd_3\mid L_3(\mathbf{x}),\ d'd_4\mid L_4(\mathbf{x})\},$$

and

$$\Lambda_3=\{\mathbf{x}\in\mathbb{Z}^2: d\mid L_1(\mathbf{x}),\ d\mid L_2(\mathbf{x}),\ d'\mid L_3(\mathbf{x}),\ d'\mid L_4(\mathbf{x})\}.$$

It therefore follows that the index of Λ_2 in Λ_3 is

$$\frac{\rho(dd_1,dd_2,d'd_3,d'd_4)}{\rho(d,d,d',d')},$$

and hence that

$$\rho_0(d_1,d_2,d_3,d_4)=\frac{\rho(dd_1,dd_2,d'd_3,d'd_4)}{\rho(d,d,d',d')}.$$

We now see that $\rho_0(p^\alpha,p^\beta,p^\gamma,p^\delta)=\rho(p^\alpha,p^\beta,p^\gamma,p^\delta)$ if $p\nmid dd'$, by the multiplicative property (3.10). It therefore follows that $E_p(d,d')=E_p$ for $p\nmid dd'$, with E_p as in Theorem 1.

We now define

$$N=\prod_{E_p=0}p$$

so that we must have $\prod_p\sigma_p(d,d')=0$ unless $N\mid dd'$. For a typical prime factor p of dd' let $p^u\,\|\,d$ and $p^v\,\|\,d'$, so that

$$\rho(d,d,d',d')=\prod_{p\mid dd'}\rho(p^u,p^u,p^v,p^v).$$

Assuming now that $N\mid dd'$ we set

$$F_N=\prod_{p\nmid N}E_p(1-\chi(p)/p)^4.$$

Moreover we define $E_p^{(u,v)}$ by (1.14), so that $E_p = E_p^{(0,0)}$. We then see that

$$\frac{\prod_p \sigma_p(d,d')}{\rho(d,d,d',d')} = F_N \prod_{p|dd'} g(p^u,p^v),$$

where

$$g(p^u,p^v) = \begin{cases} E_p^{(u,v)}(1-\chi(p)/p)^4 & \text{if } p \mid N, \\ E_p^{(u,v)}/E_p^{(0,0)} & \text{if } p \nmid 2N. \end{cases}$$

If we extend $g(m,n)$ by the multiplicativity condition

$$g(ef,e'f') = g(e,e')g(f,f') \quad \text{if } \mathrm{hcf}(ee',ff') = 1,$$

we then deduce that (6.3) holds with

$$C(d,d') = 4\pi^4 F_N g(d,d')$$

when $N \mid dd'$, and $C(d,d') = 0$ otherwise. Although we have defined $g(p^u,p^v)$ for all nonnegative integer exponents u,v the reader should note that only the values $u,v = 0,1$ are relevant for us, since d and d' may be taken to be square-free in (6.1).

When $p \nmid \Delta$ we see from (3.12) that $E_p = 1 + O(p^{-1})$ and

$$E_p^{(u,v)} \le \frac{(u+1)^2}{p^{2u}}\{1+O(p^{-1})\},$$

for $u \ge v$. Thus

$$g(p^u,p^v) \le \tau(p^u)^3\tau(p^v)^3[p^u,p^v]^{-2}$$

for $p \gg_{\Delta,N} 1$. For the remaining primes $p \ll_{\Delta,N} 1$, and in particular those primes which divide Δ, we automatically have

$$g(p^u,p^v) \ll_\Delta \tau(p^u)^3\tau(p^v)^3[p^u,p^v]^{-2}, \quad (0 \le u,v \le 1).$$

We may now deduce the required bound (6.4), with an implied constant depending on Δ, using the multiplicative property of the function $g(d,d')$.

We have now established the asymptotic formula (6.5) and the bound (6.4), and it remains to consider the constant C given by (6.6). Our work thus far shows that

$$C = \frac{\chi(\nu\nu')}{16} 4\pi^4 F_N \sum_{\substack{dd' \equiv \nu\nu' \pmod 4 \\ N|dd'}} \mu(d)\mu(d')g(d,d').$$

We shall rewrite this as

$$\frac{\pi^4 F_N}{4} \sum_{2\nmid dd',\, N\mid dd'} \frac{\chi(\nu\nu') + \chi(dd')}{2} \mu(d)\mu(d')g(d,d') = \frac{\pi^4 F_N}{8} \{\chi(\nu\nu')\Sigma_1 + \Sigma_2\},$$

where

$$\Sigma_1 = \sum_{2\nmid dd',\, N\mid dd'} \mu(d)\mu(d')g(d,d') \quad \text{and} \quad \Sigma_2 = \sum_{N\mid dd'} \chi(dd')\mu(d)\mu(d')g(d,d').$$

To evaluate Σ_1 we set $d = ef$ where $e \mid N$ and $(f,N) = 1$, and similarly $d' = e'f'$. Then

$$\Sigma_1 = \Big\{ \sum_{e,e'\mid N,\, N\mid ee'} \mu(e)\mu(e')g(e,e') \Big\} \Big\{ \sum_{(ff',2N)=1} \mu(f)\mu(f')g(f,f') \Big\},$$

so that we may use the multiplicative property to deduce that

$$\Sigma_1 = \prod_{p\mid N} \{-g(1,p) - g(p,1) + g(p,p)\} \prod_{p\nmid 2N} \{1 - g(1,p) - g(p,1) + g(p,p)\},$$

whence

$$\begin{aligned} F_N\Sigma_1 &= F_N \prod_{p\mid N} \{E_p^{(0,0)} - E_p^{(0,1)} - E_p^{(1,0)} + E_p^{(1,1)}\}(1-\chi(p)/p)^4 \\ &\qquad \times \prod_{p\nmid 2N} \{E_p^{(0,0)} - E_p^{(0,1)} - E_p^{(1,0)} + E_p^{(1,1)}\}/E_p^{(0,0)} \\ &= \prod_{p\neq 2} \{E_p^{(0,0)} - E_p^{(0,1)} - E_p^{(1,0)} + E_p^{(1,1)}\}(1-\chi(p)/p)^4, \end{aligned} \tag{7.1}$$

since $E_p^{(0,0)} = 0$ when $p \mid N$. In exactly the same way we find that

$$F_N\Sigma_2 = \prod_{p\neq 2} \{E_p^{(0,0)} - \chi(p)E_p^{(0,1)} - \chi(p)E_p^{(1,0)} + E_p^{(1,1)}\}(1-\chi(p)/p)^4. \tag{7.2}$$

Using the functions $T_\chi(p)$ and $T_\pm(p)$ given by (1.13) and (1.17) we therefore deduce that

$$C = \frac{\pi^4}{8} \left\{ \chi(\nu\nu') \prod_{p\neq 2} T_-(p)(1-\chi(p)/p)^4 + \prod_{p\neq 2} T_\chi(p)(1-\chi(p)/p)^4 \right\}.$$

This suffices for Theorem 2, providing that we can confirm the evaluation of σ_2 and σ_∞, and verify that $E_p^{(1,0)} = E_p^{(0,1)} = 0$ for any prime $p \equiv -1 \pmod 4$ that does not divide $\Delta_{12}\Delta_{34}$.

8 Proof of Theorem 2—local densities

We begin this section by defining and then computing the local densities for the variety given by (1.7), subject to the condition $\mathbf{x} \in \mathcal{R}_2$. For a prime $p > 2$ the p-adic density σ_p is merely

$$\sigma_p = \lim_{e\to\infty} p^{-4e} N(p^e), \tag{8.1}$$

where

$$N(p^e) = \#\left\{ \begin{matrix} x_1, \ldots, x_6 \\ (\text{mod } p^e) \end{matrix} : \begin{matrix} L_1(x_1,x_2)L_2(x_1,x_2) \equiv x_3^2 + x_4^2 \pmod{p^e}, \\ L_3(x_1,x_2)L_4(x_1,x_2) \equiv x_5^2 + x_6^2 \pmod{p^e} \end{matrix} \right\}.$$

Similarly, for $p = 2$ the 2-adic density in $\mathcal{R}_2$ will be given by (8.1), for $p = 2$, but with

$$N(2^e) = \#\left\{ \begin{matrix} x_1, \ldots, x_6 \\ (\text{mod } 2^e) \end{matrix} : \begin{matrix} 2 \nmid x_1, \\ L_1(x_1,x_2)L_2(x_1,x_2) \equiv x_3^2 + x_4^2 \pmod{2^e}, \\ L_3(x_1,x_2)L_4(x_1,x_2) \equiv x_5^2 + x_6^2 \pmod{2^e} \end{matrix} \right\}. \tag{8.2}$$

Finally, the real density is given by

$$\sigma_\infty = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{x_1,\ldots,x_6} e(\alpha Q_1 + \beta Q_2)dx_1 \ldots dx_6\, d\beta d\alpha,$$

where

$$Q_1 = L_1(x_1,x_2)L_2(x_1,x_2) - x_3^2 - x_4^2, \quad Q_2 = L_3(x_1,x_2)L_4(x_1,x_2) - x_5^2 - x_6^2.$$

Here (x_1, x_2) runs over $\mathcal{R}$, and x_3, x_4, x_5, x_6 each run over an interval of the form $[-cX, cX]$, with c a suitably large constant. According to part (iii) of **NC2**, this is sufficient.

For a prime $p \equiv 1 \pmod 4$ one easily finds that

$$\#\{x, y \pmod{p^e} : x^2 + y^2 \equiv A \pmod{p^e}\} = \begin{cases} p^e + ep^{e-1}(p-1) & \text{if } p^e \mid A, \\ (1+\nu_p(A))p^{e-1}(p-1) & \text{if } \nu_p(A) < e, \end{cases}$$

for any integer A, where $\nu_p(A)$ is the value of ν for which $p^\nu \,\|\, A$. Similarly, when $p \equiv -1 \pmod 4$ we have

$$\#\{x, y \pmod{p^e} : x^2 + y^2 \equiv A \pmod{p^e}\} = \begin{cases} p^{2[e/2]} & \text{if } p^e \mid A, \\ p^{e-1}(p+1) & \text{if } \nu_p(A) < e,\ 2 \mid \nu_p(A), \\ 0 & \text{if } \nu_p(A) < e,\ 2 \nmid \nu_p(A). \end{cases} \tag{8.3}$$

Finally, for $p = 2$ we have

$$\#\{x, y \pmod{2^e} : x^2 + y^2 \equiv A \pmod{2^e}\} = 2^{e+1}, \tag{8.4}$$

providing that $e \geq 2$ and $A \equiv 1 \pmod 4$.

It follows that, for a fixed prime $p \equiv 1 \pmod 4$, we have

$$N(p^e) = \sum_{x_1, x_2} p^{2e-2}(p-1)^2\{1 + \nu_p(L_1(\mathbf{x})L_2(\mathbf{x}))\}\{1 + \nu_p(L_3(\mathbf{x})L_4(\mathbf{x}))\} + O(e^2 p^{3e})$$

as $e \to \infty$, where the summation is for $\mathbf{x} \pmod{p^e}$, subject to the condition that $p^e \nmid L_1(\mathbf{x})L_2(\mathbf{x})$ and $p^e \nmid L_3(\mathbf{x})L_4(\mathbf{x})$. Now, if $\nu_1, \nu_2, \nu_3, \nu_4 < e$, then we see that

$$\begin{aligned}
&\#\{\mathbf{x} \pmod{p^e} : \nu_p(L_i(\mathbf{x})) = \nu_i,\ (1 \leq i \leq 4)\} \\
&\quad = \sum_{f_1, f_2, f_3, f_4 = 0,1} (-1)^{f_1+f_2+f_3+f_4} \#\{\mathbf{x} \pmod{p^e} : p^{\nu_i + f_i} \mid L_i(\mathbf{x}),\ (1 \leq i \leq 4)\} \\
&\quad = \sum_{f_1, f_2, f_3, f_4 = 0,1} (-1)^{f_1+f_2+f_3+f_4} p^{2e} \rho(p^{\nu_1+f_1}, p^{\nu_2+f_2}, p^{\nu_3+f_3}, p^{\nu_4+f_4})^{-1}.
\end{aligned} \tag{8.5}$$

It therefore follows that

$$\begin{aligned}
N(p^e) &= p^{4e-2}(p-1)^2 \sum_{\nu_1+\nu_2<e,\, \nu_3+\nu_4<e} (1+\nu_1+\nu_2)(1+\nu_3+\nu_4) \times \\
&\qquad \sum_{f_1, f_2, f_3, f_4 = 0,1} (-1)^{f_1+f_2+f_3+f_4} \rho(p^{\nu_1+f_1}, p^{\nu_2+f_2}, p^{\nu_3+f_3}, p^{\nu_4+f_4})^{-1} \\
&\qquad + O(e^2 p^{3e}) \\
&= p^{4e-2}(p-1)^2 \sum_{\nu_1, \nu_2, \nu_3, \nu_4 = 0}^{\infty} (1+\nu_1+\nu_2)(1+\nu_3+\nu_4) \times \\
&\qquad \sum_{f_1, f_2, f_3, f_4 = 0,1} (-1)^{f_1+f_2+f_3+f_4} \rho(p^{\nu_1+f_1}, p^{\nu_2+f_2}, p^{\nu_3+f_3}, p^{\nu_4+f_4})^{-1} \\
&\qquad + O(e^2 p^{3e}) \\
&= p^{4e-2}(p-1)^2 \sum_{\mu_1, \mu_2, \mu_3, \mu_4 = 0}^{\infty} \rho(p^{\mu_1}, p^{\mu_2}, p^{\mu_3}, p^{\mu_4})^{-1} \times \\
&\qquad \sum_{0 \leq f_i \leq \min(1, \mu_i)} (-1)^{f_1+f_2+f_3+f_4} (1+\mu_1+\mu_2-f_1-f_2)(1+\mu_3+\mu_4-f_3-f_4) \\
&\qquad + O(e^2 p^{3e}).
\end{aligned}$$

The sum over the f_i vanishes unless $\min(\mu_1, \mu_2) = \min(\mu_3, \mu_4) = 0$, in which case it is 1. We now conclude that

$$\begin{aligned}\sigma_p &= (1-1/p)^2 \sum_{\min(\mu_1,\mu_2)=\min(\mu_3,\mu_4)=0} \rho(p^{\mu_1}, p^{\mu_2}, p^{\mu_3}, p^{\mu_4})^{-1} \qquad (8.6)\\ &= (1-1/p)^2 T_-(p).\end{aligned}$$

We proceed to investigate the case $p \equiv -1 \pmod 4$ in much the same way. Using (8.3) and (8.5) we deduce that

$$N(p^e) = p^{4e-2}(p+1)^2 \sum_{\mu_1,\mu_2,\mu_3,\mu_4=0}^{\infty} (-1)^{\mu_1+\mu_2+\mu_3+\mu_4} \rho(p^{\mu_1}, p^{\mu_2}, p^{\mu_3}, p^{\mu_4})^{-1} F + O(e^2 p^{3e}),$$

where F is the number of integers f_1, f_2, f_3, f_4 in the range $0 \le f_i \le \min(1, \mu_i)$ such that $f_1 + f_2 \equiv \mu_1 + \mu_2 \pmod 2$ and $f_3 + f_4 \equiv \mu_3 + \mu_4 \pmod 2$. The sum over the f_i therefore equals 4 if $\mu_i \ge 1$ for every i, and equals 1 when $\min(\mu_1, \mu_2) = \min(\mu_3, \mu_4) = 0$. In the remaining case the sum is equal to 2. From this we deduce that

$$\sigma_p = (1+1/p)^2 T_+(p) \quad (p \equiv -1 \pmod 4)). \qquad (8.7)$$

The formula (1.12) therefore follows.

We turn next to the case of $p = 2$. In view of part (iv) of **NC2**, we will have $L_1(\mathbf{x})L_2(\mathbf{x}) \equiv L_3(\mathbf{x})L_4(\mathbf{x}) \equiv 1 \pmod 4$, providing that $2 \nmid x_1$. According to (8.2) and (8.4) we deduce that

$$N(2^e) = 2^{2e+2}\#\{\mathbf{x} \pmod{2^e} : 2 \nmid x_1\} = 2^{4e+1},$$

whence

$$\sigma_2 = 2.$$

Finally, to evaluate σ_∞, we restrict x_3, x_4, x_5, x_6 to be nonnegative, and substitute $q_1 = L_1(x_1, x_2)L_2(x_1, x_2) - x_3^2 - x_4^2$ for x_4, and similarly $q_2 = L_3(x_1, x_2)L_4(x_1, x_2) - x_5^2 - x_6^2$ for x_6. We write

$$G_1 = L_1(x_1, x_2)L_2(x_1, x_2) - x_3^2 - q_1, \quad G_2 = L_3(x_1, x_2)L_4(x_1, x_2) - x_5^2 - q_2,$$

and we set

$$F(q_1, q_2) = \frac{1}{4} \int_{x_1,x_2,x_3,x_5} G_1^{-1/2} G_2^{-1/2} dx_1 dx_2 dx_3 dx_5,$$

where the integral is subject to $(x_1, x_2) \in \mathcal{R}$ and $0 \le x_3, x_5 \le cX$, together with the constraints

$$L_1(x_1,x_2)L_2(x_1,x_2) - x_3^2 \ge q_1, \quad L_3(x_1,x_2)L_4(x_1,x_2) - x_5^2 \ge q_2.$$

Then we have

$$\sigma_\infty = 16 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{q_1,q_2} F(q_1,q_2) e(\alpha q_1 + \beta q_2) dq_1 dq_2 \, d\beta d\alpha,$$

and by the Fourier inversion theorem this reduces to $16F(0,0)$. To evaluate $F(0,0)$ we observe that

$$\int_0^{\sqrt{A}} \{A - x^2\}^{-1/2} dx = \frac{\pi}{2},$$

whence $F(0,0) = \pi^2 \operatorname{meas} \mathcal{R}/16$ and

$$\sigma_\infty = \pi^2 \operatorname{meas} \mathcal{R}.$$

Suppose next that the equations (1.7) has an integer solution $x_1, \ldots, x_6$ with $(x_1, x_2) \in \mathcal{R}_2$. It follows from part (iv) of **NC2** that $x_3^2 + x_4^2$ and $x_5^2 + x_6^2$ are nonzero integers, so that the solution is nonsingular. A standard argument now shows that this solution can be lifted via Hensel's Lemma to a positive p-adic density of points, for any prime p. Thus we must have $\sigma_p > 0$ for every p.

We now evaluate σ_p when $p \nmid \Delta$. For such primes, (3.12) gives

$$\rho(p^{\mu_1}, p^{\mu_2}, p^{\mu_3}, p^{\mu_4}) = p^{a+b}$$

where a is the maximum of the μ_i, and if $a = \mu_j$, say, then b is the maximum of the set $\{\mu_1, \mu_2, \mu_3, \mu_4\} \setminus \{\mu_j\}$. When $\min(\mu_1, \mu_2) = \min(\mu_3, \mu_4) = 0$ we therefore have

$$\rho(p^{\mu_1}, p^{\mu_2}, p^{\mu_3}, p^{\mu_4}) = p^{\mu_1+\mu_2+\mu_3+\mu_4}, \tag{8.8}$$

so that (8.6) yields

$$\begin{aligned}
\sigma_p &= (1 - 1/p)^2 \sum_{\min(\mu_1,\mu_2)=\min(\mu_3,\mu_4)=0} p^{-\mu_1-\mu_2-\mu_3-\mu_4} \\
&= (1 - 1/p)^2 \Big\{ \sum_{\min(m,n)=0} p^{-m-n} \Big\}^2 \\
&= (1 + 1/p)^2,
\end{aligned}$$

when $p \equiv 1 \pmod 4$. This proves (1.15) for such primes.

The computation for the case $p \equiv -1 \pmod 4$ is somewhat more involved. We first evaluate

$$S_1 = \sum_{\min(\mu_1,\mu_2)=0,\ \min(\mu_3,\mu_4)=0} (-1)^{\mu_1+\mu_2+\mu_3+\mu_4} \rho(p^{\mu_1}, p^{\mu_2}, p^{\mu_3}, p^{\mu_4})^{-1}.$$

Using the argument of the previous paragraph we find that

$$\begin{aligned} S_1 &= \sum_{\min(\mu_1,\mu_2)=0,\ \min(\mu_3,\mu_4)=0} (-1)^{\mu_1+\mu_2+\mu_3+\mu_4} p^{-\mu_1-\mu_2-\mu_3-\mu_4} \\ &= \Big\{ \sum_{\min(m,n)=0} (-1)^{m+n} p^{-m-n} \Big\}^2 \\ &= (p-1)^2(p+1)^{-2}. \end{aligned}$$

Next we consider

$$S_2 = \sum_{\mu_1,\mu_2,\mu_3 \geq 1} (-1)^{\mu_1+\mu_2+\mu_3} \rho(p^{\mu_1}, p^{\mu_2}, p^{\mu_3}, 1)^{-1}.$$

We may write this as

$$\begin{aligned} S_2 &= \sum_{a,b,c \geq 1} (-1)^{a+b+c} p^{\min(a,b,c)} p^{-a-b-c} \\ &= \sum_{k=1}^{\infty} p^k \sum_{\min(a,b,c)=k} (-1)^{a+b+c} p^{-a-b-c} \\ &= \sum_{k=1}^{\infty} p^k \Big\{ \sum_{a,b,c=k}^{\infty} (-1)^{a+b+c} p^{-a-b-c} - \sum_{a,b,c=k+1}^{\infty} (-1)^{a+b+c} p^{-a-b-c} \Big\} \\ &= \sum_{k=1}^{\infty} p^k \left\{ \Big(\frac{(-p^{-1})^k}{1+p^{-1}}\Big)^3 - \Big(\frac{(-p^{-1})^{k+1}}{1+p^{-1}}\Big)^3 \right\} \\ &= \frac{1+p^{-3}}{(1+p^{-1})^3} \sum_{k=1}^{\infty} p^k (-p^{-1})^{3k} \\ &= -\frac{1+p^{-3}}{(1+p^{-1})^3} \frac{1}{p^2+1}. \end{aligned}$$

Of course we get the same result for any sum in which three of the μ_i are at least 1 and the fourth is 0. The next sum to compute is

$$S_3 = \sum_{\mu_1,\mu_2 \geq 1} (-1)^{\mu_1+\mu_2} \rho(p^{\mu_1}, p^{\mu_2}, 1, 1)^{-1}.$$

This is easily found to be

$$S_3 = \sum_{a,b \geq 1} (-1)^{a+b} p^{-a-b} = (p+1)^{-2}.$$

Now if

$$S_4 = \sum_{\mu_1,\mu_2\geq 1,\ \min(\mu_3,\mu_4)=0} (-1)^{\mu_1+\mu_2+\mu_3+\mu_4}\rho(p^{\mu_1},p^{\mu_2},p^{\mu_3},p^{\mu_4})^{-1},$$

then

$$S_4 = 2S_2 + S_3 = -\frac{(1-p^{-1})^2}{(1+p^{-1})^2}\frac{1}{p^2+1}.$$

Clearly we have the same result if the rôles of μ_1,μ_2 and μ_3,μ_4 are interchanged. Finally we examine

$$S_5 = \sum_{\mu_1,\mu_2,\mu_3,\mu_4=1}^{\infty} (-1)^{\mu_1+\mu_2+\mu_3+\mu_4}\rho(p^{\mu_1},p^{\mu_2},p^{\mu_3},p^{\mu_4})^{-1}.$$

Now, according to (3.12) we have

$$\begin{aligned} S_5 &= p^{-2}\sum_{\mu_1,\mu_2,\mu_3,\mu_4=0}^{\infty} (-1)^{\mu_1+\mu_2+\mu_3+\mu_4}\rho(p^{\mu_1},p^{\mu_2},p^{\mu_3},p^{\mu_4})^{-1} \\ &= p^{-2}\{S_1+2S_4+S_5\}, \end{aligned}$$

whence

$$S_5 = \frac{S_1+2S_4}{p^2-1} = -S_4.$$

Then, as in the proof of (8.7), we have

$$\sigma_p = (1+1/p)^2\{S_1+4S_4+4S_5\} = (1+p^{-1})^2 S_1 = (1-p^{-1})^2.$$

This establishes (1.15) when $p\equiv -1 \pmod 4$.

Having dealt with the evaluation of the densities σ_p, our next task is to interpret the sums $E_p^{(u,v)}$ given by (1.14). Only primes $p\equiv -1 \pmod 4$ need concern us. We claim that whenever $p\equiv -1 \pmod 4$ we have

$$E_p^{(u,v)} = p^{-2u-2v}(1+1/p)^{-4}\lim_{e\to\infty} p^{-6e}N^{(u,v)}(p^e), \tag{8.9}$$

where

$$N^{(u,v)}(p^e) = \#\left\{ \begin{matrix} x_1,\dots,x_{10} \\ \pmod{p^e} \end{matrix} : \begin{matrix} L_1(x_1,x_2)\equiv p^u(x_3^2+x_4^2) \pmod{p^e}, \\ L_2(x_1,x_2)\equiv p^u(x_5^2+x_6^2) \pmod{p^e}, \\ L_3(x_1,x_2)\equiv p^v(x_7^2+x_8^2) \pmod{p^e}, \\ L_4(x_1,x_2)\equiv p^v(x_9^2+x_{10}^2) \pmod{p^e} \end{matrix} \right\}.$$

If $p^u \mid L_1(x_1,x_2)$, then the number of pairs x_3,x_4 modulo p^e for which

$$p^{-u}L_1(x_1,x_2)\equiv x_3^2+x_4^2 \pmod{p^{e-u}}$$

will be given by (8.3). Thus if $p^e \mid L_1(x_1, x_2)$ there are $O(p^e)$ such pairs. Otherwise suppose that $p^f \parallel L_1(x_1, x_2)$. Then if $f-u$ is even there are $p^{e+u-1}(p+1)$ pairs, and if $f-u$ is odd there are no such pairs. If we set $u_1 = u_2 = u$ and $u_3 = u_4 = v$ we then find that

$$N^{(u,v)}(p^e) = p^{4e+2u+2v-4}(p+1)^4 \sum_{\substack{0 \le \nu_i < e \\ \nu_i \equiv u_i \pmod 2}} N(p^e; \nu_1, \nu_2, \nu_3, \nu_4) + O(p^{5e}),$$

where

$$N(p^e; \nu_1, \nu_2, \nu_3, \nu_4) = \#\{x_1, x_2 \pmod{p^e} : \nu_p(L_i(x_1, x_2)) = \nu_i, \ \ (1 \le i \le 4)\}.$$

The sum over the ν_i may be re-written as

$$\sum_{\substack{0 \le \nu_i < e \\ \nu_i \equiv u_i \pmod 2}} \sum_{f_1, f_2, f_3, f_4 = 0,1} (-1)^{f_1+f_2+f_3+f_4} \frac{p^{2e}}{\rho(p^{\nu_1+f_1}, p^{\nu_2+f_2}, p^{\nu_3+f_3}, p^{\nu_4+f_4})},$$

whence

$$\lim_{e \to \infty} p^{-6e} N^{(u,v)}(p^e) = p^{2u+2v}(1+1/p)^4 \Sigma,$$

with

$$\begin{aligned} \Sigma &= \sum_{\nu_i \equiv u_i \pmod 2} \sum_{f_1, f_2, f_3, f_4 = 0,1} \frac{(-1)^{f_1+f_2+f_3+f_4}}{\rho(p^{\nu_1+f_1}, p^{\nu_2+f_2}, p^{\nu_3+f_3}, p^{\nu_4+f_4})} \\ &= \sum_{g_i \ge u_i} (-1)^{g_1+g_2+g_3+g_4} \rho(p^{g_1}, p^{g_2}, p^{g_3}, p^{g_4})^{-1} \end{aligned}$$

as in our treatment of (8.7). This suffices for the proof of (8.9).

It is now clear that $E_p^{(u,v)} \ge 0$ for $p \equiv -1 \pmod 4$. Now let $p \nmid \Delta_{12}\Delta_{34}$ for some prime $p \equiv -1 \pmod 4$, and let $u = u_1 = u_2 = 1$ and $v = u_3 = u_4 = 0$, say. Suppose we have a solution to the congruences

$$\begin{aligned} L_1(x_1, x_2) \equiv p(x_3^2 + x_4^2), \quad L_2(x_1, x_2) \equiv p(x_5^2 + x_6^2) \pmod{p^e}, \\ L_3(x_1, x_2) \equiv x_7^2 + x_8^2, \quad L_4(x_1, x_2) \equiv x_9^2 + x_{10}^2 \pmod{p^e} \end{aligned}$$

in which $p^{2f} \mid x_1, x_2$ for some exponent $2f \le e-2$. Then p^f must divide each of $x_3, \ldots, x_{10}$ and therefore

$$\begin{aligned} L_1(y_1, y_2) \equiv p(y_3^2 + y_4^2), \quad L_2(y_1, y_2) \equiv p(y_5^2 + y_6^2) \pmod{p^{e-2f}}, \\ L_3(y_1, y_2) \equiv y_7^2 + y_8^2, \quad L_4(y_1, y_2) \equiv y_9^2 + y_{10}^2 \pmod{p^{e-2f}} \end{aligned}$$

where $x_i = p^{2f} y_i$ for $i = 1, 2$ and $x_i = p^f y_i$ for $3 \le i \le 10$. Since the first two of these congruences imply that $p \mid L_1(y_1, y_2), L_2(y_1, y_2)$ we deduce that $p \mid y_1, y_2$, since $p \nmid \Delta_{12}$. It follows that $p \mid L_3(y_1, y_2), L_4(y_1, y_2)$, and hence that p divides both $y_7^2 + y_8^2$ and $y_9^2 + y_{10}^2$. Thus $p^2 \mid y_7^2 + y_8^2, y_9^2 + y_{10}^2$, so that $p^2 \mid L_3(y_1, y_2), L_4(y_1, y_2)$. Since $p \nmid \Delta_{34}$ this requires $p^2 \mid y_1, y_2$, whence, finally, $p^{2f-2} \mid x_1, x_2$. We therefore conclude that any solution of the original congruences must have $p^{e-1} \mid x_1, x_2$. In view of (8.3) we deduce that $N^{(1,0)}(p^e) = O(p^{4e})$, whence $E_p^{(1,0)} = 0$, by (8.9). Similarly we will have $E_p^{(0,1)} = 0$.

It remains to show that if $\varepsilon = -1$ then the variety (1.7) has no points with $(x_1, x_2) \in \mathcal{R}_2$. Clearly, if $\varepsilon = -1$ then we must have $T_-(p) = \pm T_+(p)$ for every prime $p \mid \Delta$ with $p \equiv -1 \pmod 4$. Let

$$\mathcal{P} = \{p \mid \Delta : p \equiv -1 \pmod 4,\ T_-(p) = -T_+(p)\}.$$

We now argue by contradiction, assuming that we have a point $(x_1, x_2) \in \mathcal{R}_2$ on the variety (1.7). Then, since $L_i(x_1, x_2) \neq 0$ by part (iv) of **NC2**, we see that the equations (1.7) entail

$$\begin{aligned} \nu_p(L_1(x_1, x_2)) &\equiv \nu_p(L_2(x_1, x_2)) \pmod 2,\\ \nu_p(L_3(x_1, x_2)) &\equiv \nu_p(L_4(x_1, x_2)) \pmod 2, \end{aligned}$$

for any prime $p \equiv -1 \pmod 4$. We now suppose that

$$2 \mid \nu_p(L_1(x_1, x_2)) - u \quad \text{and} \quad 2 \mid \nu_p(L_3(x_1, x_2)) - v$$

with $0 \leq u, v \leq 1$. Then we can find a nonsingular p-adic solution to the equations

$$\begin{aligned} L_1(x_1, x_2) &= p^u(y_3^2 + y_4^2), \quad & L_2(x_1, x_2) &= p^u(y_5^2 + y_6^2),\\ L_3(x_1, x_2) &= p^v(y_7^2 + y_8^2), \quad & L_4(x_1, x_2) &= p^v(y_9^2 + y_{10}^2). \end{aligned}$$

This can then be lifted by the standard procedure to show, via (8.9), that $E_p^{(u,v)} > 0$. Thus

$$E_p^{(u,v)} > 0 \quad \text{if} \quad 2 \mid \nu_p(L_1(x_1, x_2)) - u \quad \text{and} \quad 2 \mid \nu_p(L_3(x_1, x_2)) - v. \tag{8.10}$$

We now show that $\nu_p(L_1(x_1, x_2))$ and $\nu_p(L_3(x_1, x_2))$ have opposite parities whenever $p \in \mathcal{P}$. Since $T_-(p) = -T_+(p)$ for such a prime, and $E_p^{(u,v)} \geq 0$ for all u, v, we will have $E_p^{(0,0)} = E_p^{(1,1)} = 0$. The claim then follows from (8.10).

Conversely we now show that if $\nu_p(L_1(x_1, x_2))$ and $\nu_p(L_3(x_1, x_2))$ have opposite parities, and $p \equiv -1 \pmod 4$, then $p \in \mathcal{P}$. For such a prime, it follows from (8.10) that either $E_p^{(1,0)} > 0$ or $E_p^{(0,1)} > 0$. However we have already seen that $E_p^{(1,0)} = E_p^{(0,1)} = 0$ unless $p \mid \Delta_{12}\Delta_{34}$. Thus if $\nu_p(L_1(x_1, x_2))$ and $\nu_p(L_3(x_1, x_2))$ have opposite parities, and $p \equiv -1 \pmod 4$, then $p \mid \Delta$. Thus p must occur in the product for ε, whence $T_-(p) = \pm T_+(p)$. Since either $E_p^{(1,0)} > 0$ or $E_p^{(0,1)} > 0$ we cannot have $T_-(p) = T_+(p)$, so that we must indeed have $p \in \mathcal{P}$.

We have therefore shown that the set $\mathcal{P}$ consists precisely of those primes $p \equiv -1 \pmod 4$ which divide $L_1(x_1, x_2)L_3(x_1, x_2)$ to an odd power. Since part (iii) of **NC2** implies that $L_1(x_1, x_2)L_3(x_1, x_2)$ is positive, we conclude from part (iv) of **NC2** that

$$\chi(\nu\nu') \equiv L_1(x_1, x_2)L_3(x_1, x_2) \equiv (-1)^{\#\mathcal{P}} \pmod 4. \tag{8.11}$$

On the other hand we have

$$\prod_{p|\Delta,\, \chi(p)=-1} T_-(p)/T_+(p) = (-1)^{\#\mathcal{P}},$$

and since $\varepsilon = -1$ we deduce that

$$(-1)^{\#\mathcal{P}} = -\chi(\nu\nu').$$

This contradicts (8.11), and therefore completes the proof of Theorem 2.

9 Examples

In this section we shall discuss Theorem 2 in the context of the examples (1.10), (1.11) and (1.18). We begin with (1.10), which we repeat here as

$$y_1(y_1 + 4y_2) = x_3^2 + x_4^2, \quad (7y_1 + 16y_2)(19y_1 + 44y_2) = x_5^2 + x_6^2.$$

This has been shown to have no nontrivial rational points, even though it has nonsingular points in every completion of $\mathbb{Q}$. We take the region $\mathcal{R}^{(0)}$ to be the square $(0, 1)^2$, so that parts (i), (ii) and (iii) of **NC2** will be satisfied. Moreover part (iv) is clearly satisfied with $\nu = 1$ and $\nu' = -1$.

The existence of nonsingular local points is sufficient to ensure that $\sigma_p > 0$ for every prime p. However for the forms in (1.10) we find that $\Delta_{12}\Delta_{34} = 2^4$, so that $E_p^{(1,0)} = E_p^{(0,1)} = 0$ for any primes entering into the product in (1.16). It follows that $T_-(p) = T_+(p)$ for such primes, so that $\varepsilon = \chi(\nu\nu') = \chi(-1) = -1$. Thus the failure of the Hasse Principle is fully explained by Theorem 2, at least as far as points with $(y_1, y_2) \in \mathcal{R}_2$ are concerned.

We turn now to the example (1.11), namely

$$y_1(y_1 + 4y_2) = x_3^2 + x_4^2, \quad (y_1 + 8y_2)(13y_1 + 64y_2) = x_5^2 + x_6^2.$$

Although there are rational points in this example, we showed in §1 that all such points have $y_2/y_1 \geq -1/8$. We shall therefore consider the application of Theorem 2 to two different regions. We begin by examining the case

$$y_1, y_1 + 4y_2 > 0, \quad y_1 + 8y_2 < 0, \quad 13y_1 + 64y_2 < 0,$$

for which there are no rational points. Here we must replace L_3 and L_4 by $-L_3$ and $-L_4$ respectively, to produce linear forms which will all be positive. Having made this change we then take $\mathcal{R}^{(0)} = (0,1)^2$. Then parts (i), (ii) and (iii) of **NC2** will hold. We also see that part (iv) holds, with $\nu = 1$ and $\nu' = -1$. We may now proceed as in the previous example, noting that $\Delta_{12}\Delta_{34} = 2^5 \cdot 5$. Once again it follows that $\varepsilon = -1$, so that $\mathcal{R}_2$ produces no solutions.

On the other hand, if we look at the case

$$y_1, y_1 + 4y_2 > 0, \quad y_1 + 8y_2 > 0, \quad 13y_1 + 64y_2 > 0,$$

we may again work with $\mathcal{R}^{(0)} = (0,1)^2$. This time we have $\nu = \nu' = 1$ in part (iv) of Normalization Condition 2. The value $\Delta_{12}\Delta_{34} = 2^5 \cdot 5$ is the same as before, so that (1.16) yields $\varepsilon = \chi(\nu\nu') = \chi(1) = 1$. It therefore follows that the density of rational points in $\mathcal{R}_2$ is twice the product of local densities, while the density of rational points in the first case was of course zero.

The examples we have looked at so far all have $\varepsilon = \pm 1$. However other values may occur, as the example (1.18)

$$x_1(x_1 + 12x_2) = x_3^2 + x_4^2, \quad (x_1 + 4x_2)(x_1 + 16x_2) = x_5^2 + x_6^2,$$

will demonstrate. We shall use the region

$$\mathcal{R} = \big\{0 < x_1, x_1 + 16x_2 < X\big\}$$

so that

$$\sigma_\infty = \pi^2 \operatorname{meas} \mathcal{R} = \frac{\pi^2}{16} X^2.$$

There is a nonsingular rational point with $(x_1, x_2) = (1, 0)$, and this is enough to ensure that all the local densities are positive. Since $\Delta_{12}\Delta_{34} = 2^4 \cdot 3^2$ and $\nu = \nu' = 1$, we now find that $\varepsilon = T_-(3)/T_+(3)$. In order to show that $\varepsilon \neq \pm 1$ it will suffice to demonstrate that $E_3^{(0,0)}$ and $E_3^{(1,0)}$ are positive. To do this we shall use (8.9). When $u = v = 0$ the congruences

$$\begin{gathered} x_1 \equiv x_3^2 + x_4^2 \pmod{3}, \quad x_1 + 12x_2 \equiv x_5^2 + x_6^2 \pmod{3}, \\ x_1 + 4x_2 \equiv x_7^2 + x_8^2 \pmod{3}, \quad x_1 + 16x_2 \equiv x_9^2 + x_{10}^2 \pmod{3} \end{gathered}$$

have a nonsingular solution with $x_1 = 1$ and $x_2 = 0$, which is sufficient to ensure that $E_3^{(0,0)} > 0$. Similarly, for $u = 1$, $v = 0$, the congruences

$$\begin{gathered} x_1 \equiv 3(x_3^2 + x_4^2) \pmod{3^e}, \quad x_1 + 12x_2 \equiv 3(x_5^2 + x_6^2) \pmod{3^e}, \\ x_1 + 4x_2 \equiv x_7^2 + x_8^2 \pmod{3^e}, \quad x_1 + 16x_2 \equiv x_9^2 + x_{10}^2 \pmod{3^e} \end{gathered}$$

require $x_1 = 3x_1'$, say, so that they are equivalent to

$$x_1' \equiv x_3^2 + x_4^2 \pmod{3^{e-1}}, \quad x_1' + 4x_2 \equiv x_5^2 + x_6^2 \pmod{3^{e-1}},$$
$$3x_1 + 4x_2 \equiv x_7^2 + x_8^2 \pmod{3^e}, \quad 3x_1 + 16x_2 \equiv x_9^2 + x_{10}^2 \pmod{3^e}.$$

There is now a nonsingular solution with $x_1' = x_2 = 1$, so that $E_3^{(1,0)} > 0$, as required.

Thus (1.8) provides an example with $0 < 1 + \varepsilon < 2$. We illustrate this example numerically. Since $\sigma_2 = 2$, we see that (1.15) yields

$$\prod_p \sigma_p = \frac{2\sigma_3}{(1-1/3)^2} \prod_p (1 + \chi(p)/p)^2 = \frac{18}{\pi^2}\sigma_3.$$

Moreover one finds from (1.12) that

$$\sigma_3\Big(1 + \frac{T_-(3)}{T_+(3)}\Big) = \frac{16}{9}(T_+(3) + T_-(3)) = \frac{32}{9}(E_3^{(0,0)} + E_3^{(1,1)}).$$

One may now evaluate $E_3^{(0,0)}$ and $E_3^{(1,1)}$ by a somewhat tedious calculation along the lines of that given in the previous section to prove (1.15). The starting point is the fact that (3.12) remains true for $p = 3$, except when $\min(e_1, e_2) > \max(e_3, e_4)$, in which case

$$\rho(3^{e_1}, 3^{e_2}, 3^{e_3}, 3^{e_4}) = 3^{e_1+e_2-1},$$

or $\min(e_3, e_4) > \max(e_1, e_2)$, in which case

$$\rho(3^{e_1}, 3^{e_2}, 3^{e_3}, 3^{e_4}) = 3^{e_3+e_4-1}.$$

The conclusion is that

$$E_3^{(0,0)} = \frac{9}{20} \quad \text{and} \quad E_3^{(1,1)} = \frac{1}{20}.$$

It follows that we will have asymptotically $2X^2$ solutions to (1.18) in $\mathcal{R}_2$. This is illustrated by Table 2, in which

$$S(X) = \sum_{\mathbf{x}\in\mathcal{R}_2} r(L_1(\mathbf{x})L_2(\mathbf{x}))r(L_3(\mathbf{x})L_4(\mathbf{x})).$$

10 Acknowledgements

Parts of this work were undertaken while the author was visiting the University of Hong Kong, and the Max-Planck Institute for Mathematics in Bonn. Their hospitality and financial support is gratefully acknowledged.

The author is also extremely grateful to Professors Colliot-Thélène and Salberger for their helpful remarks on the proof of the Hasse Principle and Weak Approximation for the variety (1.7).

Table 2

X	$S(X)$	$S(X)/2X^2$
1000	1993472	0.9967 ...
2000	8030592	1.0038 ...
4000	32057728	1.0018 ...
8000	1276046726	0.9969 ...
16000	511437824	0.9989 ...
32000	2043518720	0.9978 ...

References

[1] A. Balog, Linear equations in primes, *Mathematika* **39** (1992) 367–378

[2] J.-L. Colliot-Thélène, D. Coray, and J.-J. Sansuc, Descente et principe de Hasse pour certaines variétés rationnelles, *J. reine angew. Math.* **320** (1980) 150–191

[3] J.-L. Colliot-Thélène, J.-J. Sansuc, and H.P.F. Swinnerton-Dyer, Intersections of two quadrics and Châtelet surfaces. I, *J. reine angew. Math.* **373** (1987) 37–107

[4] S. Daniel, On the divisor-sum problem for binary forms, *J. reine. angew. Math.* **507** (1999) 107–129

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

[6] V.A. Iskovskih, A counterexample to the Hasse principle for systems of two quadratic forms in five variables, *Mat. Zametki* **10** (1971) 253–257

[7] M. Nair, Multiplicative functions of polynomial values in short intervals, *Acta Arith.* **62** (1992) 257-269

D. R. Heath-Brown,
Mathematical Institute,
24-29 St Giles',
Oxford OX1 3LB, England
e-mail: rhb@maths.ox.ac.uk