ISRAEL JOURNAL OF MATHEMATICS **191** (2012), 973–1012
DOI: 10.1007/s11856-012-0019-y

# BINARY FORMS AS SUMS OF TWO SQUARES AND CHÂTELET SURFACES

BY

R. DE LA BRETÈCHE

*Institut de Mathématiques de Jussieu, Université Paris Diderot*
*UFR de Mathématiques, Case 7012, Bâtiment Chevaleret*
*75205 Paris cedex 13, France*
*e-mail: breteche@math.jussieu.fr*

AND

T. D. BROWNING

*School of Mathematics, University of Bristol*
*Bristol, BS8 1TW, United Kingdom*
*e-mail: t.d.browning@bristol.ac.uk*

ABSTRACT

The representation of integral binary forms as sums of two squares is discussed and applied to establish the Manin conjecture for certain Châtelet surfaces over $\mathbb{Q}$.

CONTENTS



Received June 30, 2010 and in revised form March 15, 2011

## 1. Introduction

Let $X$ be a proper smooth model of the affine surface

$$y^2 - az^2 = f(x), \tag{1.1}$$

where $a \in \mathbb{Z}$ is not a square and $f \in \mathbb{Z}[x]$ is a polynomial of degree 3 or 4 without repeated roots. This defines a Châtelet surface over $\mathbb{Q}$ and we will be interested here in providing a quantitative description of the density of $\mathbb{Q}$-rational points on $X$. The anticanonical linear system $|-K_X|$ has no base point and gives a morphism $\psi : X \to \mathbb{P}^4$. This paper is motivated by a conjecture of Manin [11] applied to the counting function

$$N(B) = \#\{x \in X(\mathbb{Q}) : (H_4 \circ \psi)(x) \leqslant B\},$$

for a suitably metrized exponential height $H_4 : \mathbb{P}^4(\mathbb{Q}) \to \mathbb{R}_{>0}$, whose precise definition we will delay until §5. The conjecture predicts that $N(B) \sim c_X B(\log B)^{r_X - 1}$ for some constant $c_X > 0$, where $r_X$ is the rank of the Picard group associated to $X$. Peyre [17] has given a conjectural interpretation of the constant $c_X$.

Getting an upper bound for $N(B)$ is considerably easier and the second author [5] has shown that $N(B) \ll B(\log B)^{r_X - 1}$ for any Châtelet surface. When suitable assumptions are made on $a$ and $f$ in (1.1) one can go somewhat further. Henceforth we assume that $a = -1$. In recent joint work of the authors with Peyre [4], the Manin conjecture is confirmed for a family of Châtelet surfaces that corresponds to $f(x)$ splitting completely into linear factors over $\mathbb{Q}$ in (1.1). Our aim in the present investigation is to better understand the behaviour of $N(B)$ when the factorisation of $f(x)$ into irreducibles contains an irreducible polynomial of degree 3. Here, as throughout this paper, we take irreducibility to mean irreducibility over $\mathbb{Q}$. In this case it follows from work of Colliot-Thélène, Sansuc and Swinnerton-Dyer [6, 7] that $X$ satisfies the Hasse principle and weak approximation. Moreover, it is straightforward to calculate that $r_X = 2$ (see [5, Lemma 1], for example). With this in mind we see that the following result confirms the Manin prediction.

THEOREM 1: *We have $N(B) \sim c_X B \log B$, as $B \to \infty$, where $c_X$ is the constant predicted by Peyre.*

Our result bears comparison with recent work of Iwaniec and Munshi [15], where a counting function analogous to $N(B)$ is studied as $B \to \infty$. However,

using methods based on the Selberg sieve, they are only able to produce a lower bound for the counting function which is essentially of the correct order of magnitude, a deficit that is remedied by our result.

Fix a constant $c > 0$ once and for all. We will work with compact subsets $\mathcal{R} \subset \mathbb{R}^2$ whose boundary is a piecewise continuously differentiable closed curve of length

$$\partial(\mathcal{R}) \leqslant c \sup_{\mathbf{x}=(x_1,x_2)\in\mathcal{R}} \max\{|x_1|, |x_2|\} = cr_\infty,$$

say. For any parameter $X > 0$ let $X\mathcal{R} = \{X\mathbf{x} : \mathbf{x} \in \mathcal{R}\}$. Our proof of the theorem relies upon estimating the sum

$$S(X) = \sum_{\mathbf{x}\in\mathbb{Z}^2\cap X\mathcal{R}} r\big(L(\mathbf{x})\big) r\big(C(\mathbf{x})\big),$$

where $r$ denotes the sum of two squares function, and $L, C$ are suitable binary forms of degree 1 and 3, respectively, that are defined over $\mathbb{Z}$. Recall that $r(n) = 4\sum_{d|n} \chi(d)$, where $\chi$ is the non-principal character modulo 4. For any $\mathbf{d} = (d_1, d_2) \in \mathbb{N}^2$ we let

$$\varrho(\mathbf{d}) = \varrho(\mathbf{d}; L, C) = \#\{\mathbf{x} \in \mathbb{Z}^2 \cap [0, d_1 d_2)^2 : d_1 \mid L(\mathbf{x}),\ d_2 \mid C(\mathbf{x})\}. \tag{1.2}$$

Furthermore, we define $\mathcal{E}$ to be the set of $m \in \mathbb{N}$ such that there exists $\ell \in \mathbb{Z}_{\geqslant 0}$ for which $m \equiv 2^\ell \pmod{2^{\ell+2}}$. We denote by $\mathcal{E} \pmod{2^n}$ the projection of $\mathcal{E}$ modulo $2^n$. The following result forms the technical core of this paper.

THEOREM 2: *Let $\varepsilon > 0$ and let*

$$\eta = 1 - \frac{1 + \log\log 2}{\log 2} > 0.086.$$

*Let $C \in \mathbb{Z}[\mathbf{x}]$ be an irreducible cubic form and let $L \in \mathbb{Z}[\mathbf{x}]$ be a non-zero linear form. Assume that $L(\mathbf{x}) > 0$ and $C(\mathbf{x}) > 0$ for every $\mathbf{x} \in \mathcal{R}$. Then we have*

$$S(X) = \pi^2 \operatorname{vol}(\mathcal{R}) X^2 \prod_p K_p + O\big(X^2 (\log X)^{-\eta+\varepsilon}\big),$$

*where*

$$K_p = \Big(1 - \frac{\chi(p)}{p}\Big)^2 \sum_{\nu_1,\nu_2 \geqslant 0} \frac{\chi(p^{\nu_1+\nu_2})\varrho(p^{\nu_1}, p^{\nu_2})}{p^{2\nu_1+2\nu_2}}$$

*if $p > 2$ and*

$$K_2 = 4 \lim_{n\to\infty} 2^{-2n} \#\left\{ \mathbf{x} \in (\mathbb{Z}/2^n\mathbb{Z})^2 : \begin{array}{l} L(\mathbf{x}) \in \mathcal{E} \pmod{2^n} \\ C(\mathbf{x}) \in \mathcal{E} \pmod{2^n} \end{array} \right\}.$$

*The implied constant in this estimate is allowed to depend on $\varepsilon, L, C$ and $r_\infty$.*

The sum $S(X)$ is directly linked to the density of integral points on the affine variety

$$L(\mathbf{x}) = s_1^2 + t_1^2, \quad C(\mathbf{x}) = s_2^2 + t_2^2.$$

Arguing along similar lines to the proof of [2, Theorem 4], one can interpret the leading constant in our estimate for $S(X)$ as a product of local densities for this variety. In fact this variety is related to a certain intermediate torsor that parametrises rational points on the Châtelet surfaces under consideration in this paper.

The asymptotic formula in Theorem 2 should be taken as part of an ongoing programme to understand the average order of arithmetic functions running over the values of binary quartic forms. One of the starting points for this topic lies in the work of Daniel [8], where the analogue of $S(X)$ is estimated asymptotically with $r(L)r(C)$ replaced by $r(x_1^4 + x_2^4)$. A treatment of $r(L_1) \cdots r(L_4)$ for non-proportional linear forms $L_1, \dots, L_4$ has been accomplished by Heath-Brown [12], which in turn has been improved by the authors [2]. Moreover, our allied investigation [3] could easily be adapted to handle the analogue of $S(X)$ featuring $r(L_1)r(L_2)r(Q)$ when $L_1, L_2$ are non-proportional linear forms and $Q$ is an irreducible binary quadratic form. Dealing with $r(Q_1)r(Q_2)$, for non-proportional irreducible quadratic forms $Q_1, Q_2$, or even $r(F)$ for a general irreducible quartic form $F$, seems to present a more serious challenge.

Acknowledgements. It is pleasure to thank the referee for carefully reading the manuscript and making numerous helpful comments, including drawing our attention to an oversight in the original treatment of Lemma 11. While working on this paper the first author was supported by Institute de Mathématiques de Jussiea and the second author was supported by EPSRC grant number EP/E053262/1. Part of this work was carried out while the second author was visiting the first author at the Université Paris 7 Denis Diderot, funded by ANR "Points entiers points rationnels".

## 2. Polynomials modulo $n$

Our analysis will require information about the number of solutions to various systems of polynomial equations modulo $n$. For any polynomial $f \in \mathbb{Z}[x]$ of degree $d \geqslant 2$, we define the content of $f$ to be the greatest common divisor of its coefficients. Thus a polynomial has content 1 if and only if it is primitive.

Let

$$\varrho_f(n) = \#\{x \in \mathbb{Z}/n\mathbb{Z} : f(x) \equiv 0 \,(\mathrm{mod}\, n)\}. \tag{2.1}$$

Since $\varrho_f(n)$ is a multiplicative function of $n$ it will suffice to analyse it for prime powers. We begin by recording the following upper bounds.

LEMMA 1: *Assume that* $\mathrm{disc}(f) \neq 0$ *and that* $p$ *is a prime which does not divide the content of* $f$*, with* $p^\mu \| \mathrm{disc}(f)$*. Then for any* $\nu \geqslant 1$ *we have*

$$\varrho_f(p^\nu) \leqslant d \min\left\{p^{\frac{\mu}{2}}, p^{(1-\frac{1}{d})\nu}, p^{\nu-1}\right\}.$$

*Proof.* The inequality $\varrho_f(p^\nu) \leqslant dp^{\frac{\mu}{2}}$ is due to Huxley [14] and the inequality $\varrho_f(p^\nu) \leqslant dp^{(1-\frac{1}{d})\nu}$ is due to Stewart [18, Corollary 2]. The final inequality is trivial. ∎

One of the ingredients in our work will be the Dedekind zeta function

$$\zeta_k(s) = \sum_{\mathfrak{a}} \frac{1}{N_{k/\mathbb{Q}}(\mathfrak{a})^s} = \prod_{\mathfrak{p}} \Big(1 - \frac{1}{N_{k/\mathbb{Q}}(\mathfrak{p})^s}\Big)^{-1},$$

for $\Re e(s) > 1$, when $k$ is a number field obtained by adjoining to $\mathbb{Q}$ the root of an irreducible polynomial $f \in \mathbb{Z}[x]$. Here $\mathfrak{a}$ runs over the set of non-zero integral ideals in $k$ and $\mathfrak{p}$ runs over prime ideals. By a well-known principle due to Dedekind [10, p. 212], for a rational prime $p \nmid f_0 \,\mathrm{disc}(f)$, where $f_0$ denotes the leading coefficient of $f$, we have the ideal factorisation $(p) = \mathfrak{p}_1^{e_1}\mathfrak{p}_2^{e_2}\cdots$, with $N_{k/\mathbb{Q}}(\mathfrak{p}_i) = p^{r_i}$, corresponding to the factorisation

$$f(x) \equiv f_1(x)^{e_1} f_2(x)^{e_2} \cdots \,(\mathrm{mod}\, p)$$

for polynomials $f_i(x)$ of degree $r_i$ which are irreducible modulo $p$. When $r_i = 1$ the polynomial $f_i$ has a root modulo $p$. Thus, for $p \nmid f_0 \,\mathrm{disc}(f)$, we have

$$\varrho_f(p) = \#\{\mathfrak{p} : N_{k/\mathbb{Q}}(\mathfrak{p}) = p\}.$$

The Eulerian factors of $\zeta_k(s)$ which correspond to prime ideals $\mathfrak{p}$ for which $N_{k/\mathbb{Q}}(\mathfrak{p}) = p^r$ for $r \geqslant 2$, or $p \mid f_0 \,\mathrm{disc}(f)$, define a holomorphic and bounded function in the half-plane $\Re e(s) > \frac{1}{2}$, without any zeros there.

We will need to investigate the Dirichlet series

$$G_f(s) = \sum_{n=1}^{\infty} \frac{\varrho_f(n)}{n^s}, \quad G_f(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)\varrho_f(n)}{n^s}, \tag{2.2}$$

for $\Re e(s) > 1$, where $\chi$ is the real non-principal character modulo 4. Let $\kappa \in (0, \frac{1}{d})$. It follows from Lemma 1 that for any $p \mid f_0 \operatorname{disc}(f)$ we have

$$\sum_{\nu \geqslant 1} \frac{\varrho_f(p^\nu)}{p^{\nu(1-\kappa)}} \ll_\kappa 1.$$

Hence for all $\kappa \in (0, \frac{1}{d})$ there exists an arithmetic function $h$ such that

$$G_f(s) = \zeta_k(s) \sum_{n=1}^{\infty} \frac{h(n)}{n^s} = \zeta_k(s) H_f(s),$$

say, with $\sum_{n=1}^{\infty} |h(n)| n^{-1+\kappa} \ll_\kappa 1$. In the same manner $G_f(s, \chi)$ is related to the Hecke $L$-function

$$L(s, \chi) = \sum_{\mathfrak{a}} \frac{\chi(N_{k/\mathbb{Q}}(\mathfrak{a}))}{N_{k/\mathbb{Q}}(\mathfrak{a})^s},$$

defined for $\Re e(s) > 1$. Note that when $d$ is odd $L(s, \chi)$ will be analytic at $s = 1$, since $\chi$ is a quadratic character. Thus we have $G_f(s, \chi) = L(s, \chi) H_f(s, \chi)$, where

$$H_f(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n) h(n)}{n^s}.$$

The following result is well-known and follows on combining the above with the results contained in the survey of Heilbronn [13].

LEMMA 2: *Let $A > 0$ and let $f \in \mathbb{Z}[x]$ be an irreducible cubic polynomial with content 1. Then we have*

$$\sum_{n \leqslant X} \frac{\chi(n) \varrho_f(n)}{n} = \vartheta(f; \chi) + O_A\big((\log X)^{-A}\big),$$

*with $\vartheta(f; \chi) = L(1, \chi) H_f(1, \chi)$. Furthermore, we have*

$$\sum_{p \leqslant X} \frac{\chi(p) \varrho_f(p)}{p} \ll 1.$$

In the present investigation we will be concerned with the case $f(x) = C(x, 1)$, an irreducible polynomial of degree $d = 3$ defined over $\mathbb{Z}$. We will need to relate the series

$$D(s) = \sum_{n=1}^{\infty} \frac{\chi(n) \varrho(1, n)}{n^{1+s}} \tag{2.3}$$

to $G_{C(x,1)}(s,\chi)$, where $\varrho(d_1,d_2)$ is given by (1.2). To this end it will be necessary to have some further information about the size of $\varrho(d_1,d_2)$ at prime powers. We will suppose once and for all that

$$L(\mathbf{x}) = ax_1 + bx_2, \quad C(\mathbf{x}) = c_0x_1^3 + c_1x_1^2x_2 + c_2x_1x_2^2 + c_3x_2^3, \tag{2.4}$$

for $a, b, c_i \in \mathbb{Z}$, with non-zero integers

$$\Delta = |\operatorname{Res}(L,C)|, \quad \Delta' = |\operatorname{disc}(C)|. \tag{2.5}$$

Our investigation is summarised in the following result.

LEMMA 3: *Let $C \in \mathbb{Z}[\mathbf{x}]$ be an irreducible cubic form and let $L \in \mathbb{Z}[\mathbf{x}]$ be a non-zero linear form. Assume that $L, C$ are primitive and let $\Delta$, $\Delta'$ be as in* (2.5). *Then we have the following expressions.*

(1) *When $p \nmid c_0\Delta'$ and $\nu \in \mathbb{N}$ then we have*

$$\varrho(1,p^\nu) = \begin{cases} p^{\nu-1}(p^{[\frac{\nu}{3}]}-1)\varrho_{C(x,1)}(p) + p^{\nu+[\frac{\nu}{3}]} & \text{if } \nu \equiv 0 \,(\mathrm{mod}\, 3), \\ p^{\nu-1}(p^{[\frac{\nu}{3}]+1}-1)\varrho_{C(x,1)}(p) + p^{\nu+[\frac{\nu}{3}]-1} & \text{if } \nu \equiv 1 \,(\mathrm{mod}\, 3), \\ p^{\nu-1}(p^{[\frac{\nu}{3}]+1}-1)\varrho_{C(x,1)}(p) + p^{\nu+[\frac{\nu}{3}]} & \text{if } \nu \equiv 2 \,(\mathrm{mod}\, 3). \end{cases}$$

*In particular, when $p \nmid c_0\Delta'$ we have*

$$\varrho(1,p) = (p-1)\varrho_{C(x,1)}(p) + 1.$$

*For any prime $p$ and $\nu \in \mathbb{N}$, we have*

$$\varrho(1,p^\nu) \ll \min\{p^{2\nu-1}, p^{\frac{4\nu}{3}}\}.$$

(2) *When $\nu_2 \leqslant 3\nu_1$ and $p \nmid \Delta$, we have*

$$\varrho(p^{\nu_1},p^{\nu_2}) \leqslant p^{\nu_1+2\nu_2-\lceil\frac{\nu_2}{3}\rceil}.$$

*When $0 \leqslant 3\nu_1 < \nu_2$ and $p \nmid c_0\Delta\Delta'$, we have*

$$\varrho(p^{\nu_1},p^{\nu_2}) \leqslant \Big(3+\frac{1}{p}\Big)p^{2\nu_1+\nu_2+[\frac{\nu_2}{3}]}.$$

(3) *For any prime $p$ and $\nu_1, \nu_2 \in \mathbb{Z}_{\geqslant 0}$ we have*

$$\varrho(p^{\nu_1},p^{\nu_2}) \ll \min\{p^{\nu_1+2\nu_2}, p^{2\nu_1+2\nu_2-1}, p^{2\nu_1+\frac{4\nu_2}{3}}\}.$$

*Proof.* These expressions are founded on a preliminary study of the related quantity

$$\varrho^*(p^{\nu_1},p^{\nu_2}) = \#\{\mathbf{x} \in \mathbb{Z}^2 \cap [0,p^{\nu_1+\nu_2})^2 : p^{\nu_1} \mid L(\mathbf{x}),\ p^{\nu_2} \mid C(\mathbf{x}),\ p \nmid \mathbf{x}\}. \tag{2.6}$$

We will follow the convention that $\varrho^*(1,1)=1$. We can relate this quantity to $\varrho(p^{\nu_1},p^{\nu_2})$ via the easily checked identity

$$\varrho(p^{\nu_1},p^{\nu_2})=\sum_{0\leqslant k\leqslant \max\{\nu_1,\lceil\frac{\nu_2}{3}\rceil\}}\varrho^*\big(p^{\max\{\nu_1-k,0\}},p^{\max\{\nu_2-3k,0\}}\big)p^{m_k}, \tag{2.7}$$

with $m_k=2(\min\{\nu_1,k\}+\min\{\nu_2,3k\}-k)$. This follows on partitioning the $\mathbf{x}$ to be counted according to the common $p$-adic order of $x_1,x_2$ and $p^{\max\{\nu_1,\lceil\frac{\nu_2}{3}\rceil\}}$.

Proceeding with our analysis of $\varrho^*(p^{\nu_1},p^{\nu_2})$, we begin by noting that

$$\varrho^*(1,p^\nu)=\varphi(p^\nu)\varrho_{C(x,1)}(p^\nu) \tag{2.8}$$

if $p\nmid c_0$, since the solutions $\mathbf{x}$ to be counted satisfy $p\nmid x_2$ for $p\nmid c_0$. Hence Lemma 1 yields $\varrho^*(1,p^\nu)\leqslant 3\varphi(p^\nu)$ if $p\nmid c_0\Delta'$. Suppose now that $p\mid c_0\Delta'$. If $\mathbf{x}$ is counted by $\varrho^*(1,p^\nu)$ then $\xi\leqslant v_p(c_0)$ if $p^\xi\| x_2$. We conclude from Lemma 1 that

$$\varrho^*(1,p^\nu)\leqslant\sum_{0\leqslant\xi\leqslant v_p(c_0)}\varphi(p^{\nu-\xi})\cdot p^\xi\varrho_{p^{-\xi}C(x,p^\xi)}(p^{\nu-\xi})\ll p^\nu, \tag{2.9}$$

where we recall our convention that the implied constants are allowed to depend on the coefficients of $L,C$. This latter estimate holds for any prime $p$. Next we note that

$$\varrho^*(p^{\nu_1},p^{\nu_2})\leqslant\min\{p^{2\nu_2}\varrho^*(p^{\nu_1},1),p^{2\nu_1}\varrho^*(1,p^{\nu_2})\}.$$

Since $\varrho^*(p^{\nu_1},p^{\nu_2})=0$ when $\min\{\nu_1,\nu_2\}>v_p(\Delta)$, and $\varrho^*(p^{\nu_1},1)=\varphi(p^{\nu_1})$, it therefore follows from (2.9) that

$$\varrho^*(p^{\nu_1},p^{\nu_2})\ll p^{\nu_1+\nu_2}. \tag{2.10}$$

We are now ready to deduce the statement of Lemma 3. When $p\nmid\Delta'$ and $\nu\geqslant 1$ it follows from Hensel's lemma that $\varrho_{C(x,1)}(p^\nu)=\varrho_{C(x,1)}(p)$. The first pair of displayed relations in part (1) now follow directly from (2.7) and (2.8). The final part is again based on (2.7), but now combined with (2.9).

Turning to the proof of part (2), for which we call upon (2.7), we see that when $\nu_2\leqslant 3\nu_1$ and $p\nmid\Delta$ we have

$$\begin{aligned}\varrho(p^{\nu_1},p^{\nu_2})&=\sum_{\lceil\frac{\nu_2}{3}\rceil\leqslant k\leqslant\nu_1}p^{2\nu_2}\varrho^*(p^{\nu_1-k},1)\\&=p^{2\nu_2}\sum_{\lceil\frac{\nu_2}{3}\rceil\leqslant k\leqslant\nu_1}\varphi(p^{\nu_1-k})\leqslant p^{\nu_1+2\nu_2-\lceil\frac{\nu_2}{3}\rceil}.\end{aligned}$$

When $3\nu_1 < \nu_2$ and $p \nmid c_0\Delta\Delta'$ we have

$$\begin{aligned}\varrho(p^{\nu_1}, p^{\nu_2}) &= \sum_{\nu_1 \leqslant k \leqslant [\frac{\nu_2}{3}]} p^{2\nu_1+4k}\varrho^*(1, p^{\nu_2-3k}) + \Big(\Big\lceil \frac{\nu_2}{3}\Big\rceil - \Big[\frac{\nu_2}{3}\Big]\Big)p^{2\nu_1+2\nu_2-2\lceil\frac{\nu_2}{3}\rceil} \\ &\leqslant 3p^{2\nu_1+\nu_2+[\frac{\nu_2}{3}]} + \Big(\Big\lceil \frac{\nu_2}{3}\Big\rceil - \Big[\frac{\nu_2}{3}\Big]\Big)p^{2\nu_1+2\nu_2-2\lceil\frac{\nu_2}{3}\rceil} \\ &\leqslant \Big(3 + \frac{1}{p}\Big)p^{2\nu_1+\nu_2+[\frac{\nu_2}{3}]}.\end{aligned}$$

Finally, part (3) is a consequence of the inequalities

$$\varrho(p^{\nu_1}, p^{\nu_2}) \leqslant p^{2\nu_2}\varrho(p^{\nu_1}, 1) = p^{\nu_1+2\nu_2}, \quad \varrho(p^{\nu_1}, p^{\nu_2}) \leqslant p^{2\nu_1}\varrho(1, p^{\nu_2}),$$

together with part (1) of the lemma. ∎

In general, the forms $L, C$ need not be primitive. We let $\ell_1, \ell_2 \in \mathbb{N}$ and $L^*, C^*$ be primitive forms such that

$$L = \ell_1 L^*, \quad C = \ell_2 C^*.$$

One can easily restrict attention to primitive forms in Lemma 3 via the trivial observation that

$$\frac{\varrho(\mathbf{d}; L, C)}{(d_1 d_2)^2} = \frac{\varrho(\mathbf{d}'; L^*, C^*)}{(d_1' d_2')^2}, \tag{2.11}$$

for any $\mathbf{d} \in \mathbb{N}^2$, where $d_i' = \gcd(d_i, \ell_i)^{-1} d_i$.

Returning to the Dirichlet series $D(s)$ defined in (2.3), we write

$$D(s) = G_{C(x,1)}(s, \chi)A(s), \tag{2.12}$$

where $G_{C(x,1)}(s, \chi)$ is given by (2.2) and $A(s)$ is the Dirichlet series associated to an appropriate arithmetic function $a$. We will need the following result.

LEMMA 4: *For any $\varepsilon > 0$ and $\sigma \geqslant \frac{5}{6} + \varepsilon$ we have $\sum_{n=1}^{\infty} |a(n)| n^{-\sigma} \ll 1$.*

*Proof.* Since the two functions involved are multiplicative it suffices to analyse the Euler products

$$D(s) = \prod_p D_p(s), \quad G_{C(x,1)}(s, \chi) = \prod_p G_{p,C(x,1)}(s, \chi).$$

Assume $\Re e(s) = \sigma > \frac{2}{3}$. When $p \nmid c_0\Delta'$, Lemma 1 and part (1) of Lemma 3 yield

$$\begin{aligned} D_p(s) &= 1 + \frac{\chi(p)\varrho_{C(x,1)}(p)}{p^s} + O\big(p^{-2\sigma+\frac{2}{3}} + p^{-1-\sigma}\big) \\ &= G_{p,C(x,1)}(s,\chi)\Big(1 + O\big(p^{-2\sigma+\frac{2}{3}} + p^{-1-\sigma} + p^{-2\sigma}\big)\Big). \end{aligned}$$

When $p \mid c_0\Delta'$, we have

$$D_p(s) = 1 + O\big(p^{\frac{2}{3}-\sigma}\big), \quad G_{p,C(x,1)}(s,\chi) = 1 + O\big(p^{\frac{2}{3}-\sigma}\big).$$

From this we deduce that (2.12) holds with the Dirichlet series $A$ associated to a function $a$ satisfying the bound recorded in the lemma. ∎

We close this section with a simple result concerning the estimation of summatory functions that involve the convolution of arithmetic functions.

LEMMA 5: *Let $A > 0$. Let $g, h$ be arithmetic functions and $C, C', C''$ constants such that*

$$\sum_{d=1}^{\infty} \frac{|h(d)|(\log 2d)^A}{d} \leqslant C'', \quad \sum_{d\leqslant x} \frac{g(d)}{d} = C + O\Big(\frac{C'}{(\log 2x)^A}\Big).$$

*Then we have*

$$\sum_{n\leqslant x} \frac{(g*h)(n)}{n} = C\sum_{d=1}^{\infty} \frac{h(d)}{d} + O\Big(\frac{C''(C+C')}{(\log 2x)^A}\Big).$$

*Proof.* We clearly have

$$\sum_{n\leqslant x} \frac{(g*h)(n)}{n} = \sum_{d\leqslant x} \frac{h(d)}{d} \sum_{m\leqslant \frac{x}{d}} \frac{g(m)}{m}.$$

We approximate the inner sum over $m$ by $C$ if $d \leqslant \sqrt{x}$. On noting that

$$\sum_{d>\sqrt{x}} \frac{|h(d)|}{d} \leqslant \sum_{d=1}^{\infty} \frac{|h(d)|}{d} \frac{(\log 2d)^A}{(\log 2\sqrt{x})^A} \ll \frac{C''}{(\log 2x)^A},$$

we are easily led to the conclusion of the lemma. ∎

## 3. Preliminary steps

In this section we shall begin the proof of Theorem 2. Recall the notation (2.4) and (2.5) concerning $L, C$. We will find it convenient to estimate the corresponding sum $S_0(X)$, say, in which we insist that the greatest common divisor of $x_1, x_2$ is odd. Note that $r(2n) = r(n)$ for any positive integer $n$. We may therefore write

$$S(X) = \sum_{k_0 \geqslant 0} \sum_{\substack{\mathbf{x} \in \mathbb{Z}^2 \cap X\mathcal{R} \\ 2^{k_0} \| \mathbf{x}}} r(L(\mathbf{x}))r(C(\mathbf{x})) = \sum_{k_0 \geqslant 0} S_0(2^{-k_0}X).$$

We will also need to extract 2-adic factors from $L(\mathbf{x})$ and $C(\mathbf{x})$. Thus we have

$$S(X) = \sum_{k_0 \geqslant 0} \sum_{\mathbf{k}=(k_1,k_2) \in \mathbb{Z}^2_{\geqslant 0}} S_{\mathbf{k}}(2^{-k_0}X),$$

where $S_{\mathbf{k}}(X)$ is the restriction of $S(X)$ to $\mathbf{x}$ for which $2^{-k_1}L(\mathbf{x}) \equiv 1 \pmod 4$ and $2^{-k_2}C(\mathbf{x}) \equiv 1 \pmod 4$, with $2 \nmid \mathbf{x}$. In particular, it is clear that $k_1, k_2 \ll \log X$ and $\min\{k_1, k_2\} \leqslant v_2(\Delta)$ in order for $S_{\mathbf{k}}(2^{-k_0}X)$ to be non-zero. We will need to show that the available range for $k_1, k_2$ can be reduced with an acceptable error. A straightforward application of [1, Corollary 1] yields

$$S_{\mathbf{k}}(X) \ll 2^{\varepsilon(k_1+k_2)}(2^{-\max\{k_1,k_2\}}X^2 + X^{1+\varepsilon}),$$

for any $\varepsilon > 0$. It follows that

$$S(X) = \sum_{k_0 \geqslant 0} \sum_{0 \leqslant k_1, k_2 \leqslant \log\log X} S_{\mathbf{k}}(2^{-k_0}X) + O\big(X^2(\log X)^{-(1-\varepsilon)\log 2}\big). \tag{3.1}$$

The condition $2^{-k_1}L(\mathbf{x}) \equiv 1 \pmod 4$ is easy to analyse. Without loss of generality we may assume that $a$ is odd. Let $0 \leqslant c < 2^{k_1+2}$ be such that $ac \equiv -b \pmod{2^{k_1+2}}$ and $c' \in \{-1, 1\}$ such that $c' \equiv a \pmod 4$. Then we see that $2^{-k_1}L(\mathbf{x}) \equiv 1 \pmod 4$ is equivalent to the existence of $x_1' \equiv 1 \pmod 4$ such that

$$x_1 = cx_2 + c'2^{k_1}x_1'.$$

If $k_1 \geqslant 1$, the condition that $2 \nmid \mathbf{x}$ reduces to the condition that $x_2$ should be odd. If $k_1 = 0$, the condition $2 \nmid \mathbf{x}$ holds automatically.

Next we note that the condition $2^{-k_2}C(\mathbf{x}) \equiv 1 \pmod 4$ can be written

$$C(cx_2 + c'2^{k_1}x_1', x_2) \equiv 2^{k_2}{x_1'}^3 \pmod{2^{k_2+2}}.$$

If the form $C(cY + c'2^{k_1}X, Y)$ has all coefficients divisible by $2^{k_2+1}$ then this congruence has no solutions. Otherwise define $k_1' \leqslant k_2$ so that $2^{k_1'}$ is the largest power of 2 dividing all the coefficients, and set

$$C(cY + c'2^{k_1}X, Y) = 2^{k_1'}C_0(X, Y).$$

Writing $k_2' = k_2 - k_1' \geqslant 0$ then we see that the above congruence is equivalent to $C_0(x_1', x_2) \equiv 2^{k_2'}{x_1'}^3 \pmod{2^{k_2'+2}}$. Since $x_1'$ is odd we have $x_2 \equiv \alpha x_1' \pmod{2^{k_2'+2}}$, for $\alpha \in [0, 2^{k_2'+2})$ being one of the roots of

$$C_0(1, \alpha) \equiv 2^{k_2'} \pmod{2^{k_2'+2}}. \tag{3.2}$$

The condition that $x_2$ be odd, which should be added when $k_1 \geqslant 1$, is therefore equivalent to the condition that $\alpha$ be odd. Finally, we make the change of variables $x_2 = \alpha x_1' + 2^{k_2'+2}x_2'$ and note that $x_1', x_2' \ll X$ whenever $\mathbf{x} \in X\mathcal{R}$. We denote by $n(k_1, k_2)$ the number of available $\alpha$ and recall from above that $\min\{k_1, k_2\} \leqslant v_2(\Delta)$. Since $a$ is odd we clearly have

$$n(k_1, k_2) \ll \#\{x \pmod{2^{k_1+k_2}} : \ x \equiv -ba^{-1} \pmod{2^{k_1}}, \ C(x, 1) \equiv 0 \pmod{2^{k_2}}\}.$$

If $k_2 \leqslant k_1$ then the right-hand side is at most $2^{k_2} \ll 1$. If $k_2 > k_1$ then the right-hand side is at most $2^{k_1}\varrho_{C(x,1)}(2^{k_2}) \ll 1$ by Lemma 1. Hence we have

$$n(k_1, k_2) \ll 1. \tag{3.3}$$

In summary we have shown that the conditions $v_2(L(\mathbf{x})) = k_1$, $v_2(C(\mathbf{x})) = k_2$ and $2 \nmid \mathbf{x}$, with $2^{-k_1}L(\mathbf{x}) \equiv 1 \pmod 4$ and $2^{-k_2}C(\mathbf{x}) \equiv 1 \pmod 4$, can be written $\mathbf{x} = \mathbf{M}\mathbf{x}'$ with $x_1' \equiv 1 \pmod 4$ and

$$\mathbf{M} = \mathbf{M}_\alpha = \begin{pmatrix} c'2^{k_1} & c \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \alpha & 2^{k_2'+2} \end{pmatrix} = \begin{pmatrix} c'2^{k_1} + c\alpha & c2^{k_2'+2} \\ \alpha & 2^{k_2'+2} \end{pmatrix},$$

where $\alpha$ is a zero of (3.2) that should be odd when $k_1 \geqslant 1$. We note that

$$|\det \mathbf{M}| = 2^{k_1+k_2'+2}. \tag{3.4}$$

Furthermore, a little thought reveals that

$$K_2 = \sum_{k_0 \geqslant 0} \frac{1}{2^{2k_0}} \sum_{k_1, k_2 \geqslant 0} \frac{n(k_1, k_2)}{2^{k_1+k_2'+2}} = \frac{1}{3} \sum_{k_1, k_2 \geqslant 0} \frac{n(k_1, k_2)}{2^{k_1+k_2'}}, \tag{3.5}$$

in the notation of Theorem 2.

We are now ready to start our analysis of $S(X)$ in earnest, for which we follow the line of attack in [2] and [12]. In the present investigation we will not seek

complete uniformity in $L, C$ and $\mathcal{R}$, unlike in [2], which will greatly streamline our exposition. Let us set $Y = X^{\frac{1}{2}}(\log X)^{-C}$ with $C$ a large unspecified constant. When $0 < n \ll X^3$ and $n' = 2^{-v_2(n)}n \equiv 1 \pmod 4$, we write

$$\begin{aligned} r(n) = r(n') &= 4 \sum_{\substack{d_2 \mid n' \\ d_2 \leqslant X^{\frac{3}{2}}}} \chi(d_2) + 4 \sum_{\substack{e_2 \mid n' \\ e_2 > X^{\frac{3}{2}}}} \chi(e_2) \\ &= 4 \sum_{\substack{d_2 \mid n \\ d_2 \leqslant X^{\frac{3}{2}}}} \chi(d_2) + 4 \sum_{\substack{d_2 \mid n \\ n' > d_2 X^{\frac{3}{2}}}} \chi(d_2) \\ &= 4A_+(n) + 4A_-(n). \end{aligned}$$

We will apply this with $n = C(\mathbf{x})$. In the same manner, when $0 < m \ll X$ we can write

$$r(m) = 4B_+(m) + 4B_0(m) + 4B_-(m),$$

under the hypothesis that $m' = 2^{-v_2(m)}m \equiv 1 \pmod 4$, with

$$B_+(m) = \sum_{\substack{d_1 \mid m \\ d_1 \leqslant Y}} \chi(d_1), \quad B_0(m) = \sum_{\substack{d_1 \mid m \\ Y < d_1 \leqslant \frac{X}{Y}}} \chi(d_1), \quad B_-(m) = \sum_{\substack{d_1 \mid m \\ m' > d_1 \frac{X}{Y}}} \chi(d_1).$$

Making the transformation $\mathbf{x} = \mathbf{M}\mathbf{x}'$, it follows that

$$S_{\mathbf{k}}(X) = \sum_{\alpha} S_{\mathbf{k},\alpha}(X),$$

where

$$S_{\mathbf{k},\alpha}(X) = \sum_{\substack{\mathbf{x}' \in \mathbb{Z}^2 \cap X\mathcal{R}_{\mathbf{M}} \\ x_1' \equiv 1 \,(\mathrm{mod}\, 4)}} r(L_{\mathbf{M}}(\mathbf{x}'))r(C_{\mathbf{M}}(\mathbf{x}')),$$

with

$$\mathcal{R}_{\mathbf{M}} = \{\mathbf{x}' \in \mathbb{R}^2 : \mathbf{M}\mathbf{x}' \in \mathcal{R}\}, \quad L_{\mathbf{M}}(\mathbf{x}') = L(\mathbf{M}\mathbf{x}'), \quad C_{\mathbf{M}}(\mathbf{x}') = C(\mathbf{M}\mathbf{x}').$$

The region $\mathcal{R}_{\mathbf{M}}$ has volume $2^{-k_1-k_2'-2}\,\mathrm{vol}(\mathcal{R})$ and is contained in a box with side length $\ll |\det \mathbf{M}|^{-1}2^{k_1+k_2'} \ll 1$. Collecting together the above we may conclude that

$$S_{\mathbf{k}}(X) = 16 \sum_{\alpha} \sum_{\pm,\pm} S_{\pm,\pm}(X; \mathbf{k}, \alpha) + 4T(X; \mathbf{k}, \alpha), \tag{3.6}$$

with

$$S_{\pm,\pm}(X;\mathbf{k},\alpha)=\sum_{\substack{\mathbf{x}'\in\mathbb{Z}^2\cap X\mathcal{R}_{\mathbf{M}}\\ x_1'\equiv 1 \pmod 4}} A_{\pm}(C_{\mathbf{M}}(\mathbf{x}'))B_{\pm}(L_{\mathbf{M}}(\mathbf{x}')) \tag{3.7}$$

and

$$T(X;\mathbf{k},\alpha)=\sum_{\substack{\mathbf{x}'\in\mathbb{Z}^2\cap X\mathcal{R}_{\mathbf{M}}\\ x_1'\equiv 1 \pmod 4}} r(C_{\mathbf{M}}(\mathbf{x}'))B_0(L_{\mathbf{M}}(\mathbf{x}')).$$

The sums $S_{\pm,\pm}(2^{-k_0}X;\mathbf{k},\alpha)$ will make up the main term in our final asymptotic formula and we save their analysis for the following section. We dedicate the remainder of this section to showing that $T(2^{-k_0}X;\mathbf{k},\alpha)$ makes a satisfactory overall contribution

$$\sum_{k_0\geqslant 0}\sum_{0\leqslant k_1,k_2\leqslant \log\log X}\sum_{\alpha} T(2^{-k_0}X;\mathbf{k},\alpha)=T(X),$$

say, to the error term. By (3.3) we have

$$T(X)\ll(\log\log X)^2\sum_{k_0\geqslant 0}\sum_{m\in\mathcal{B}} T_m(2^{-k_0}X)|B_0(m)|,$$

where $\mathcal{B}$ is defined to be the intersection

$$\{m\in\mathbb{Z}:\exists d\mid m \text{ s.t. } Y<d\leqslant XY^{-1}\}\cap\{m\in\mathbb{Z}:\exists\mathbf{x}\in X\mathcal{R} \text{ s.t. } L(\mathbf{x})=m\}$$

and

$$T_m(X)=\sum_{\substack{\mathbf{x}\in\mathbb{Z}^2\cap X\mathcal{R}\\ L(\mathbf{x})=m}} r(C(\mathbf{x})).$$

But then [2, Lemma 6] yields

$$T(X)\ll X\frac{(\log\log X)^{\frac{17}{4}}}{(\log X)^{\eta}}\sum_{k_0\geqslant 0}\max_{m\in\mathbb{N}}|T_m(2^{-k_0}X)|,$$

where

$$\eta=1-\frac{1+\log\log 2}{\log 2}.$$

Once combined with the following result this is therefore enough to conclude the proof that $T(X)\ll X^2(\log X)^{-\eta+\varepsilon}$, which suffices for Theorem 2.

LEMMA 6: *Let $\varepsilon>0$ and let $m\leqslant X$. Then we have*

$$T_m(X)\ll X(\log X)^{\varepsilon}.$$

*Proof.* We consider here the case $a \neq 0$, the case $b \neq 0$ being dealt with similarly. The relation $L(\mathbf{x}) = m$ allows us to write $x_1 = a^{-1}(m - bx_2)$ and

$$C(\mathbf{x}) = \frac{1}{a^3} C(m - bx_2, ax_2) = \frac{1}{a^3}(c_3' x_2^3 + c_2' m x_2^2 + c_1' m^2 x_2 + c_0' m^3),$$

with

$$c_3' = C(-b, a), \quad c_2' = 3b^2 c_0 - 2abc_1 + a^2 c_2, \quad c_1' = -3bc_0 + c_1 a, \quad c_0' = c_0.$$

Let $\delta_m = \gcd_{0 \leqslant i \leqslant 3}(c_i' m^{3-i})$, so that $C_m(x_2) = a^3 \delta_m^{-1} C(\mathbf{x})$ is primitive as a polynomial in $x_2$. It follows that

$$T_m(X) \leqslant \sum_{\substack{\mathbf{x} \in \mathbb{Z}^2 \cap X\mathcal{R} \\ L(\mathbf{x}) = m}} r(a^4 C(\mathbf{x})) \leqslant \sum_{x_2 \ll X} r(a\delta_m C_m(x_2)).$$

The rest of the proof has much in common with the proof of [2, Lemma 5] and so we shall attempt to be brief.

Write $r_0(n) = \frac{1}{4} r(n)$ and $r_1$ for the multiplicative function defined via

$$r_1(p^\nu) = \begin{cases} \nu + 1, & \text{if } p \mid 3a\delta_m, \\ r_0(p^\nu), & \text{otherwise.} \end{cases}$$

We obtain

$$T_m(X) \leqslant 4\tau(a\delta_m) \sum_{x_2 \ll X} r_1(C_m(x_2)).$$

Clearly $\delta_m \mid c_3' \neq 0$, whence $\tau(a\delta_m) \ll 1$. The polynomial $C_m \in \mathbb{Z}[x_2]$ has degree 3 and is both primitive and irreducible over $\mathbb{Q}$. Therefore, the only possible fixed prime divisors are 2 and 3. An application of [1, Lemma 5] allows one to deduce that there exists $\alpha \mid 36$, $m_2, m_3 \leqslant 4$ and $\gamma = 2^{m_2} 3^{m_3}$ such that the polynomial

$$g_{\alpha,\beta}(x_2) = \frac{C_m(\alpha x_2 + \beta)}{\gamma}$$

is without any fixed prime divisor for each $\beta$ modulo $\alpha$. We obtain

$$\sum_{x_2 \ll X} r_1(C_m(x_2)) \ll \sum_{\alpha} \sum_{\beta \,(\mathrm{mod}\, \alpha)} \sum_{x_2 \ll X} r_1(g_{\alpha,\beta}(x_2)).$$

Since $\|g_{\alpha,\beta}\| \ll \|C_m\| \ll m^3$, it now follows from [1, Theorem 2] that

$$\sum_{x_2 \ll X} r_1(C_m(x_2)) \ll X \sum_{\alpha} \sum_{\beta \,(\mathrm{mod}\, \alpha)} \prod_{p \ll X} \left\{ \Big(1 - \frac{\varrho_{g_{\alpha,\beta}}(p)}{p}\Big) \sum_{\nu \geqslant 0} \frac{\varrho_{g_{\alpha,\beta}}(p^\nu) r_1(p^\nu)}{p^\nu} \right\},$$

because $X \gg m^{\varepsilon}$, where $\varrho_{g_{\alpha,\beta}}(p)$ is given by (2.1). A straightforward consideration of discriminants (see [1, Lemma 1], for example) yields $\mathrm{disc}(g_{\alpha,\beta}) \ll m^6$.

To go further it is clear that we will need good upper bounds for the function $\varrho_{g_{\alpha,\beta}}(p^\nu)$ for prime powers $p^\nu$. Such estimates are furnished by Lemma 1. Thus for any prime $p$ we deduce that

$$\sum_{\nu \geqslant 1} \frac{\varrho_{g_{\alpha,\beta}}(p^\nu) r_1(p^\nu)}{p^\nu} \ll \frac{1}{p}.$$

By including a factor

$$\ll \prod_{p \mid \mathrm{disc}(g_{\alpha,\beta})} \Big(1 + \frac{1}{p}\Big)^{O(1)} \ll (\log\log m)^{O(1)} \ll (\log X)^{\varepsilon},$$

we take care of the primes $p \mid \mathrm{disc}(g_{\alpha,\beta})$. Next, for any $p \nmid \mathrm{disc}(g_{\alpha,\beta})$, we have

$$\sum_{\nu \geqslant 2} \frac{\varrho_{g_{\alpha,\beta}}(p^\nu) r_1(p^\nu)}{p^\nu} \ll \frac{1}{p^2},$$

which allows us to ignore the exponents $\nu \geqslant 2$.

For any prime $p \geqslant 5$, we have $\varrho_{g_{\alpha,\beta}}(p) = \varrho_{C_m}(p)$, which for $p \nmid ac_3'$ is equal to $\varrho_{C(m-bx_2,ax_2)}(p)$. If $p \nmid ma$ then the map $\mathbb{Z}/p\mathbb{Z} \setminus \{mb^{-1}\} \to \mathbb{Z}/p\mathbb{Z}$, given by $x_2 \mapsto ax_2(m - bx_2)^{-1}$, is injective. It follows that $\varrho_{g_{\alpha,\beta}}(p) = \varrho_{C(1,x)}(p)$, for $p \geqslant 5$ and $p \nmid mac_3'$. Observing that $r_0(p) = 1 + \chi(p)$, our investigation so far has therefore shown that

$$\begin{aligned}\sum_{x_2 \ll X} r_1\big(g_{\alpha,\beta}(x_2)\big) &\ll X(\log X)^{\varepsilon} \prod_{\substack{p \ll X \\ p \nmid \mathrm{disc}(g_{\alpha,\beta})}} \Big(1 + \frac{\varrho_{C(1,x)}(p)(r_0(p) - 1)}{p}\Big) \\ &\ll X(\log X)^{\varepsilon} \prod_{p \ll X} \Big(1 + \frac{\chi(p)\varrho_{C(1,x)}(p)}{p}\Big) \\ &\ll X(\log X)^{\varepsilon},\end{aligned}$$

by Lemma 2. This therefore completes the proof of the lemma. ∎

## 4. Level of distribution

The focus of this section is upon estimating the sums in (3.7). For any $\mathbf{d} \in \mathbb{N}^2$ let

$$\Lambda(\mathbf{d}) = \Lambda(\mathbf{d}; L, C) = \{\mathbf{x} \in \mathbb{Z}^2 : d_1 \mid L(\mathbf{x}),\ d_2 \mid C(\mathbf{x})\}$$

and let $\Lambda_{\mathbf{M}}(\mathbf{d}) = \Lambda(\mathbf{d}; L_{\mathbf{M}}, C_{\mathbf{M}})$. Given any region $\mathcal{A} \subset \mathbb{R}^2$, we will write $X\mathcal{A}_4$ for the set $\{\mathbf{x} \in \mathbb{Z}^2 \cap X\mathcal{A} : x_1 \equiv 1 \,(\mathrm{mod}\, 4)\}$. We clearly have

$$S_{\pm,\pm}(X; \mathbf{k}, \alpha) = \sum_{\substack{d_1 \ll Y \\ d_2 \ll X^{\frac{3}{2}}}} \chi(d_1 d_2) \#(\Lambda_{\mathbf{M}}(\mathbf{d}) \cap X\mathcal{R}_4^{\pm,\pm}(\mathbf{d}, \mathbf{M})),$$

with, for example,

$$X\mathcal{R}^{-,-}(\mathbf{d}, \mathbf{M}) = \{\mathbf{x}' \in X\mathcal{R}_{\mathbf{M}} : C_{\mathbf{M}}(\mathbf{x}') > d_2 X^{\frac{3}{2}},\ L_{\mathbf{M}}(\mathbf{x}') > d_1 X Y^{-1}\}.$$

Let $\|\mathbf{M}\|$ denote the maximum modulus of any entry in the matrix $\mathbf{M}$ and let $\varrho_{\mathbf{M}}(\mathbf{d}) = \varrho(\mathbf{d}; L_{\mathbf{M}}, C_{\mathbf{M}})$, in the notation of (1.2). Loosely speaking, the idea is now to rewrite the inner cardinality as a sum of cardinalities, each one over lattice points belonging to an appropriate region. We would like to approximate each such cardinality by its volume. In doing so we need to show that the associated error term makes a satisfactory overall contribution once summed over the remaining parameters. This is the essential content of the following "level of distribution" result.

LEMMA 7: *Let $\varepsilon > 0$ and let $V_1, V_2, X \geqslant 2$. Assume that $C \in \mathbb{Z}[\mathbf{x}]$ is an irreducible cubic form and let $L \in \mathbb{Z}[\mathbf{x}]$ be a non-zero linear form. Then there exists an absolute constant $A > 0$ such that*

$$\sum_{\substack{\mathbf{d} \in \mathbb{N}^2 \\ d_i \leqslant V_i \\ 2 \nmid d_1 d_2}} \sup_{\partial(\mathcal{A}) \leqslant M} \left| \#\big(\Lambda_{\mathbf{M}}(\mathbf{d}) \cap X\mathcal{A}_4\big) - \frac{\mathrm{vol}(\mathcal{A}) X^2 \varrho_{\mathbf{M}}(\mathbf{d})}{4(d_1 d_2)^2} \right| \ll \|\mathbf{M}\|^{\varepsilon} (MX(\sqrt{V_1 V_2} + V_1) + V_1 V_2)(\log V_1 V_2)^A,$$

*where the supremum is taken over compact subsets $\mathcal{A} \subset \mathbb{R}^2$ whose boundary is a piecewise continuously differentiable closed curve with length $\partial(\mathcal{A}) \leqslant M$ and throughout which $L(\mathbf{x}) > 0$ and $C(\mathbf{x}) > 0$.*

We will not prove this result here, following closely as it does the arguments developed in [3, Lemme 5], [8, Lemma 3.2] and [16, Proposition 1]. Now it follows from (3.4) that $d_1 d_2$ is coprime to $\det \mathbf{M}$, so that $\varrho_{\mathbf{M}}(\mathbf{d}) = \varrho(\mathbf{d}; L, C) = \varrho(\mathbf{d})$. We may therefore conclude from Lemma 7 that

$$S_{\pm,\pm}(X; \mathbf{k}, \alpha) = \sum_{\substack{d_1 \ll Y \\ d_2 \ll X^{\frac{3}{2}}}} \frac{\chi(d_1 d_2)\, \mathrm{vol}(\mathcal{R}^{\pm,\pm}(\mathbf{d}, \mathbf{M})) X^2 \varrho(\mathbf{d})}{4(d_1 d_2)^2} + O\Big(\frac{2^{\varepsilon(k_1+k_2)} X^2}{(\log X)^{\frac{C}{2}-A}}\Big).$$

Choosing $C = 2A + 8$ and replacing $X$ by $2^{-k_0}X$, we see that the overall contribution from this error term is

$$\ll \sum_{k_0 \geqslant 0} \frac{(2^{-k_0}X)^2}{(\log X)^4} \sum_{k_1,k_2 \leqslant \log\log X} 2^{\varepsilon(k_1+k_2)} n(k_1,k_2) \ll \frac{X^2}{(\log X)^2},$$

by (3.3). This is satisfactory for Theorem 2.

Our final task is to produce an asymptotic formula for the sum

$$S(V_1,V_2) = \sum_{\substack{\mathbf{d}\in\mathbb{N}^2 \\ d_i \leqslant V_i}} \frac{\chi(d_1d_2)\varrho(\mathbf{d})}{(d_1d_2)^2}.$$

Recall the definition of $K_p$ from the statement of Theorem 2. We will establish the following result.

LEMMA 8: *Let $\varepsilon > 0$ and $A > 0$. For any $V_1, V_2 \geqslant 2$ we have*

$$S(V_1,V_2) = \frac{\pi^2}{16}K' + O\Big(\frac{\log V_{\min}}{(\log V_{\max})^A} + \frac{1}{(\log V_{\min})^A}\Big)$$

*where $V_{\min} = \min\{V_1,V_2\}$, $V_{\max} = \max\{V_1,V_2\}$ and $K' = \prod_{p>2} K_p$.*

*Proof.* We begin by establishing the lemma for the case in which $L$ and $C$ are both primitive. We first consider the case $V_1 \geqslant V_2$. The sum to be estimated can be written

$$S(V_1,V_2) = \sum_{d_2 \leqslant V_2} \frac{\chi(d_2)\varrho(1,d_2)}{d_2^2} S_1(V_1,d_2),$$

with

$$S_1(V_1,d_2) = \sum_{d_1 \leqslant V_1} \frac{\chi(d_1)\varrho(d_1,d_2)}{\varrho(1,d_2)d_1^2}.$$

This summand is a multiplicative arithmetic function in $d_1$ and so the associated Dirichlet series $F_1(s)$ has an Euler product $\prod_p F_{1,p}(s)$. When $p^{\nu_2} \| d_2$, we have

$$F_{1,p}(s) = \sum_{\nu_1 \geqslant 0} \frac{\chi(p^{\nu_1})\varrho(p^{\nu_1},p^{\nu_2})}{\varrho(1,p^{\nu_2})p^{\nu_1(2+s)}}.$$

In particular, when $p \nmid d_2$ we have

$$F_{1,p}(s) = \Big(1 - \frac{\chi(p)}{p^{1+s}}\Big)^{-1}$$

since $\varrho(d_1,1) = d_1$. We may therefore write $F_1(s) = L(1+s,\chi)J_1(1+s;d_2)$, where $L(1+s,\chi)$ is the Dirichlet $L$-function associated to $\chi$ and $J_1(s;d_2)$ is

the Dirichlet series associated to an arithmetic function $j_{d_2}$, with $J_1$ absolutely convergent in the half-plane $\Re e(s) \geqslant 0$. We observe that

$$J_{1,p}(1;d_2) = \Big(1 - \frac{\chi(p)}{p}\Big) F_{1,p}(0). \tag{4.1}$$

Let us write $J_1^*(s;d_2)$ for the Dirichlet series associated to $|j_{d_2}|$. For any $A > 0$, Lemma 5 yields

$$S_1(V_1,d_2) = L(1,\chi) J_1(1;d_2) + O\Big(\frac{J_1^*(\frac{3}{4};d_2)}{(\log V_1)^A}\Big).$$

Now it is clear that

$$J_1^*\Big(\frac{3}{4};d_2\Big) = \prod_{p^{\nu_2} \| d_2} J_{1,p}^*\Big(\frac{3}{4};p^{\nu_2}\Big),$$

with

$$\varrho(1,p^{\nu_2}) J_{1,p}^*\Big(\frac{3}{4};p^{\nu_2}\Big) \leqslant (1+p^{-\frac{3}{4}}) \sum_{\nu_1 \geqslant 0} \frac{\varrho(p^{\nu_1},p^{\nu_2})}{p^{\frac{7\nu_1}{4}}}.$$

We apply the inequalities in Lemma 3 to estimate $\varrho(p^{\nu_1},p^{\nu_2})$.

Suppose first that $p \nmid c_0 \Delta\Delta'$. Then $\varrho(1,p^{\nu_2}) \leqslant 4p^{\nu_2+[\frac{\nu_2}{3}]}$,

$$\sum_{1 \leqslant \nu_1 < \lceil \frac{\nu_2}{3} \rceil} \frac{\varrho(p^{\nu_1},p^{\nu_2})}{p^{\frac{7\nu_1}{4}}} \leqslant \Big(3+\frac{1}{p}\Big) \sum_{1 \leqslant \nu_1 < \lceil \frac{\nu_2}{3} \rceil} p^{\frac{\nu_1}{4}+\nu_2+[\frac{\nu_2}{3}]} \leqslant \Big(3+\frac{1}{p}\Big)\Big[\frac{\nu_2}{3}\Big] p^{\nu_2+\frac{5}{4}[\frac{\nu_2}{3}]},$$

and

$$\sum_{\nu_1 \geqslant \lceil \frac{\nu_2}{3} \rceil} \frac{\varrho(p^{\nu_1},p^{\nu_2})}{p^{\frac{7\nu_1}{4}}} \leqslant \sum_{\nu_1 \geqslant \lceil \frac{\nu_2}{3} \rceil} p^{2\nu_2 - \lceil \frac{\nu_2}{3} \rceil - \frac{3\nu_1}{4}} = \frac{p^{2\nu_2 - \frac{7}{4}\lceil \frac{\nu_2}{3} \rceil}}{1-p^{-\frac{3}{4}}}.$$

Thus

$$\begin{aligned} \frac{\varrho(1,p^{\nu_2})(J_{1,p}^*(\frac{3}{4};p^{\nu_2})-1)}{p^{\nu_2}} &\leqslant \Big(\frac{p^{\nu_2-\frac{7}{4}\lceil\frac{\nu_2}{3}\rceil}}{1-p^{-\frac{3}{4}}} + \Big(3+\frac{1}{p}\Big)\Big[\frac{\nu_2}{3}\Big] p^{\frac{5}{4}[\frac{\nu_2}{3}]}\Big)(1+p^{-\frac{3}{4}}) \\ &\quad + 4p^{[\frac{\nu_2}{3}]-\frac{3}{4}} \\ &\ll (1+\nu_2) p^{\frac{5\nu_2}{12}}. \end{aligned} \tag{4.2}$$

Suppose now that $p \mid \gcd(d_2, c_0\Delta\Delta')$. On the one hand we have

$$\sum_{\nu_1 \geqslant 0} \frac{\varrho(p^{\nu_1},p^{\nu_2})}{p^{\frac{7\nu_1}{4}}} \ll \varrho(1,p^{\nu_2}) + \sum_{\nu_1 \geqslant 1} \frac{p^{\nu_1+2\nu_2}}{p^{\frac{7\nu_1}{4}}} \ll p^{2\nu_2-\frac{3}{4}},$$

which will suffice for small values of $\nu_2$. On the other hand we have

$$\sum_{\nu_1\geqslant 0}\frac{\varrho(p^{\nu_1},p^{\nu_2})}{p^{\frac{7\nu_1}{4}}}\ll\sum_{\nu_1\leqslant\frac{2\nu_2}{3}}\frac{p^{2\nu_1+\frac{4\nu_2}{3}}}{p^{\frac{7\nu_1}{4}}}+\sum_{\nu_1>\frac{2\nu_2}{3}}\frac{p^{\nu_1+2\nu_2}}{p^{\frac{7\nu_1}{4}}}\ll p^{\frac{3\nu_2}{2}}.$$

Observe that

$$\prod_{p|c_0\Delta\Delta'}\Big(1+O\Big(\sum_{\nu_2\geqslant 1}\min\{p^{-\frac{3}{4}},p^{-\frac{\nu_2}{2}}\}\Big)\Big)\leqslant\prod_{p|c_0\Delta\Delta'}\big(1+O(p^{-\frac{3}{4}})\big),$$

which is $O(1)$. Using Dirichlet convolution these estimates allow us to conclude that

$$\sum_{d_2\leqslant V_2}\frac{\varrho(1,d_2)J_1^*(\frac{3}{4};d_2)}{d_2^2}\ll\sum_{d_2\leqslant V_2}\frac{\varrho(1,d_2)}{d_2^2}\ll\log V_2,$$

whence

$$S(V_1,V_2)=\frac{\pi}{4}\sum_{d_2\leqslant V_2}\frac{\chi(d_2)\varrho(1,d_2)J_1(1;d_2)}{d_2^2}+O\Big(\frac{\log V_2}{(\log V_1)^A}\Big).$$

The function $J_1(1;d_2)$ is multiplicative in $d_2$. Let $p\nmid c_0\Delta\Delta'$. We have

$$|J_{1,p}(1;p^{\nu_2})-1|\leqslant J_{1,p}^*(1;p^{\nu_2})-1\leqslant J_{1,p}^*\Big(\frac{3}{4};p^{\nu_2}\Big)-1.$$

Combining (4.1) with (4.2) allows us to show that for $1\leqslant\nu_2\leqslant 3$ we have

$$\varrho(1,p^{\nu_2})J_1(1;p^{\nu_2})=\varrho(1,p^{\nu_2})+O(p^{2\nu_2-\frac{7}{4}\lceil\frac{\nu_2}{3}\rceil})$$

and for $\nu_2\geqslant 4$ we have

$$\varrho(1,p^{\nu_2})J_1(1;p^{\nu_2})=\varrho(1,p^{\nu_2})+O\big((1+\nu_2)p^{\frac{5\nu_2}{12}}\big).$$

Thus, in terms of Dirichlet convolution, the function $\chi(d_2)\varrho(1,d_2)J_1(1;d_2)d_2^{-1}$ is close to $\chi(d_2)\varrho(1,d_2)d_2^{-1}$ and so to $\chi(d_2)\varrho_{C(x,1)}(d_2)$. It now follows from Lemmas 2, 4 and 5 that

$$S(V_1,V_2)=\frac{\pi}{4}\vartheta(C(x,1);\chi)K_1'+O\Big(\frac{\log V_2}{(\log V_1)^A}+\frac{1}{(\log V_2)^A}\Big),$$

for any $A > 0$, with

$$\begin{aligned} K_1' &= \vartheta(C(x,1);\chi)^{-1} \sum_{d_2 \geqslant 1} \frac{\chi(d_2)\varrho(1,d_2)J_1(1;d_2)}{d_2^2} \\ &= \prod_p \left( \frac{1-\chi(p)p^{-1}}{H_{p,C(x,1)}(1)} \sum_{\nu_2 \geqslant 0} \frac{\chi(p^{\nu_2})\varrho(1,p^{\nu_2})J_1(1;p^{\nu_2})}{p^{2\nu_2}} \right) \\ &= \prod_p \left( \frac{(1-\chi(p)p^{-1})^2}{H_{p,C(x,1)}(1)} \sum_{\nu_2 \geqslant 0} \frac{\chi(p^{\nu_2})}{p^{2\nu_2}} \sum_{\nu_1 \geqslant 0} \frac{\chi(p^{\nu_1})\varrho(p^{\nu_1},p^{\nu_2})}{p^{2\nu_1}} \right) \\ &= \frac{\pi K'}{4\vartheta(C(x,1);\chi)}. \end{aligned}$$

Here we have used (4.1) for the penultimate equality. This completes the proof of the lemma in the case $V_1 \geqslant V_2$.

Next we suppose that $V_2 \geqslant V_1$. The estimation of $S(V_1, V_2)$ in this case is completely analogous to the case we have just dealt with apart from a number of minor technical complications. We begin with the expressions

$$S(V_1,V_2) = \sum_{d_1 \leqslant V_1} \frac{\chi(d_1)\varrho(d_1,1)}{d_1^2} S_2(V_2,d_1), \quad S_2(V_2,d_1) = \sum_{d_2 \leqslant V_2} \frac{\chi(d_2)\varrho(d_1,d_2)}{\varrho(d_1,1)d_2^2}.$$

One sees that the sum $S_2(V_2, d_1)$ again involves a multiplicative arithmetic function with associated Dirichlet series $F_2(s) = \prod_p F_{2,p}(s)$. When $p \nmid d_1$, we have

$$F_{2,p}(s) = \sum_{\nu_2 \geqslant 0} \frac{\chi(p^{\nu_2})\varrho(1,p^{\nu_2})}{p^{\nu_2(2+s)}} = D_p(1+s) = G_{p,C(x,1)}(1+s)A_p(1+s),$$

where $D_p(s)$, $G_{p,C(x,1)}(s)$, $A_p(s)$ are the Eulerian factors of the Dirichlet series appearing in (2.12). When $p^{\nu_1} \| d_1$ and $p \nmid c_0\Delta\Delta'$ it follows from part (2) of Lemma 3 and the identity $\varrho(p^\nu, 1) = p^\nu$ that

$$|F_{2,p}(s) - 1| \leqslant \sum_{\nu_2 \geqslant 1} \frac{\varrho(p^{\nu_1},p^{\nu_2})}{\varrho(p^{\nu_1},1)p^{\nu_2(2+\sigma)}} \ll p^{-\frac{3}{4}},$$

for $\Re e(s) = \sigma \geqslant -\frac{1}{4}$. When $p^{\nu_1} \| d_1$ and $p \mid c_0\Delta\Delta'$ we deduce from part (3) of Lemma 3 that

$$F_{2,p}(s) \ll p^{\frac{3\nu_1}{8}},$$

for $\Re e(s) \geqslant -\frac{1}{4}$. We may therefore write $F_2(s) = G_{C(x,1)}(1+s,\chi)J_2(1+s;d_1)$ with $G_{C(x,1)}(s,\chi)$ given in (2.2) and $J_2(s;d_1)$ the Dirichlet series associated to

an arithmetic function $j_{d_1}$ which is absolutely convergent in the the half-plane $\Re e(s) > \frac{5}{6}$.

Lemmas 2, 4 and 5 now yield

$$\begin{aligned} &S(V_1, V_2) \\ &\quad = \vartheta(C(x,1);\chi) \sum_{d_1 \leqslant V_1} \frac{\chi(d_1)\varrho(d_1,1)J_2(1;d_1)}{d_1^2} + O\Big(\frac{1}{(\log V_2)^A} \sum_{d_1 \leqslant V_1} \frac{g(d_1)}{d_1}\Big), \end{aligned}$$

with $g$ a multiplicative function satisfying

$$g(p^\nu) = \begin{cases} 1 + O(p^{-\frac{3}{4}}), & \text{if } p \nmid c_0 \Delta\Delta', \\ O(p^{\frac{3\nu}{8}}), & \text{otherwise.} \end{cases}$$

This implies that

$$S(V_1, V_2) = \vartheta(C(x,1);\chi) \sum_{d_1 \leqslant V_1} \frac{\chi(d_1)J_2(1;d_1)}{d_1} + O\Big(\frac{\log V_1}{(\log V_2)^A}\Big).$$

An application of Lemma 5 yields

$$S(V_1, V_2) = \vartheta(C(x,1);\chi)\frac{\pi}{4}K_2' + O\Big(\frac{\log V_1}{(\log V_2)^A} + \frac{1}{(\log V_1)^A}\Big),$$

with

$$\begin{aligned} K_2' &= \frac{4}{\pi} \sum_{d_1 \in \mathbb{N}} \frac{\chi(d_1)J_2(1;d_1)}{d_1} \\ &= \prod_p \Big(1 - \frac{\chi(p)}{p}\Big) \sum_{\nu_1 \geqslant 0} \frac{\chi(p^{\nu_1})\varrho(p^{\nu_1},1)J_2(1;p^{\nu_1})}{p^{2\nu_1}} \\ &= \prod_p \Big(\frac{1 - \chi(p)p^{-1}}{G_{p,C(x,1)}(1,\chi)}\Big) \sum_{\nu_1 \geqslant 0} \frac{\chi(p^{\nu_1})}{p^{2\nu_1}} \sum_{\nu_2 \geqslant 0} \frac{\chi(p^{\nu_2})\varrho(p^{\nu_1},p^{\nu_2})}{p^{2\nu_2}} \\ &= \frac{\pi K'}{4\vartheta(C(x,1);\chi)}. \end{aligned}$$

This completes the proof of the lemma in the remaining case $V_2 \geqslant V_1$.

It remains to say a few words about the case in which $L, C$ are not primitive. Suppose that $L = \ell_1 L^*$ and $C = \ell_2 C^*$ for primitive forms $L^*$ and $C^*$. Then it follows from (2.11) that

$$S(V_1, V_2) = \sum_{h_i | \ell_i} \chi(h_1 h_2) S_{\frac{\ell_1}{h_1}, \frac{\ell_2}{h_2}}\Big(\frac{V_1}{h_1}, \frac{V_2}{h_2}\Big),$$

where the inner sum now involves $L^*, C^*$ and for any $\mathbf{a} \in \mathbb{N}^2$ we denote by $S_{\mathbf{a}}(V_1, V_2)$ the corresponding sum in which $\gcd(d_i, a_i) = 1$ in the summation over $\mathbf{d}$. In our case $\ell_1$ and $\ell_2$ may be viewed as absolute constants. Tracing through the argument above we are easily led to an estimate for $S_{\mathbf{a}}(V_1, V_2)$ that generalises the case $a_1 = a_2 = 1$ that we have already handled. Once inserted into the above this therefore suffices to handle the case in which $L$ or $C$ is not primitive. ∎

Combining Lemma 8 with partial summation gives

$$S_{\pm,\pm}(X; \mathbf{k}, \alpha) = X^2 \operatorname{vol}(\mathcal{R}) \frac{\pi^2 K'}{2^{8+k_1+k_2'}} + O\Big(\frac{X^2}{(\log X)^4}\Big).$$

Bringing everything together in (3.1) and (3.6) we may now conclude that

$$S(X) = \pi^2 K \operatorname{vol}(\mathcal{R}) X^2 + O\big(X^2 (\log X)^{-\eta+\varepsilon}\big),$$

with

$$K = K' \sum_{(k_0,k_1,k_2) \in \mathbb{Z}^3_{\geqslant 0}} \frac{n(k_1, k_2)}{2^{2k_0+k_1+k_2'+2}} = K' K_2,$$

by (3.5). This completes the proof of Theorem 2.

## 5. Passage to the intermediate torsors

We are now ready to commence our proof of Theorem 1. Recall the assumption in (1.1) that $a = -1$ and $f$ has degree 3 or 4, with an irreducible cubic factor without repeated roots. Thus $x_2^4 f(\frac{x_1}{x_2}) = L(\mathbf{x})C(\mathbf{x})$ with $L$ of degree 1 and $C$ of degree 3. We suppose that $L, C$ take the shape (2.4), for appropriate $a, b, c_i \in \mathbb{Z}$. Let $\delta = \sqrt{5 \max\{|a|, |b|, |c_i|\}}$. Then we will work with the norm

$$\|\mathbf{x}\| = \max\{|x_0|, |x_1|, |x_2|, \delta^{-1}|x_3|, \delta^{-1}|x_4|\},$$

in the definition of the exponential height function $H_4$ on $\mathbb{P}^4(\mathbb{Q})$.

In what follows it will be convenient to use the notation $Z^m$ for the set of primitive vectors in $\mathbb{Z}^m$. Our starting point is [5, Lemma 2], which reveals that

$$N(B) = \frac{1}{4} \#\Big\{(y, z, t; u, v) \in Z^3 \times Z^2 : \begin{array}{l} \|(v^2 t, uvt, u^2 t, y, z)\| \leqslant B, \\ y^2 + z^2 = t^2 L(u, v) C(u, v) \end{array}\Big\}.$$

We denote by $\mathcal{T} \subset \mathbb{A}^5 = \operatorname{Spec} \mathbb{Q}[y, z, t, u, v]$ the subvariety defined by the equation

$$y^2 + z^2 = t^2 L(u, v) C(u, v), \tag{5.1}$$

together with $(y,z,t) \neq \mathbf{0}$ and $(u,v) \neq \mathbf{0}$. Then $\mathcal{T}$ is a $\mathbb{G}_m^2$-torsor over $X$. We have $\|(v^2t, uvt, u^2t, y, z)\| = \max\{u^2, v^2\}|t|$, by our choice of norm function, for any $(y,z,t;u,v)$ under consideration. Since there is no solution with $t=0$ we have

$$N(B) = \frac{1}{2}\#\Big\{(y,z,t;u,v) \in (Z^3 \times Z^2) \cap \mathcal{T} : \ 0 < \max\{u^2,v^2\}t \leqslant B\Big\}. \tag{5.2}$$

The overall contribution that arises from $(y,z,t;u,v)$ for which $L(u,v)C(u,v)$ is zero is clearly $O(1)$, which is satisfactory.

Let

$$\mathfrak{D} = \{d \in \mathbb{N} : \ p \mid d \Rightarrow p \equiv 1 \,(\mathrm{mod}\, 4)\} \tag{5.3}$$

and note that $d_0 \in \mathfrak{D}$ for any $d_0 \mid d$ with $d \in \mathfrak{D}$. For $m, n \in \mathbb{N}$ we let

$$r(n;m) = \#\{a,b \in \mathbb{Z} : n = a^2 + b^2, \ \gcd(m,a,b) = 1\}.$$

Then $r(n;1) = r(n)$ is the usual $r$-function and $r(y^2n;y) = 0$ unless $y \in \mathfrak{D}$. Using the Möbius function to detect the coprimality condition we obtain

$$r(y^2n;y) = \sum_{\substack{k \mid y \\ k \in \mathfrak{D}}} \mu(k) r\Big(\frac{y^2n}{k^2}\Big),$$

for any $y \in \mathfrak{D}$. Given any $\varepsilon_1, \varepsilon_2 \in \{\pm 1\}$ and $T \geqslant 1$ we define the region

$$R^{\varepsilon_1,\varepsilon_2}(T) = \Big\{(u,v) \in \mathbb{R}^2 : \ \begin{array}{l} |u|, |v| \leqslant \sqrt{T}, \\ \varepsilon_1 L(u,v) > 0, \ \varepsilon_2 C(u,v) > 0 \end{array} \Big\}.$$

Applying the above it now follows that

$$N(B) = \frac{1}{2}\sum_{k \in \mathfrak{D}} \mu(k) \sum_{\substack{t \leqslant \frac{B}{k} \\ t \in \mathfrak{D}}} \sum_{\substack{\varepsilon_1,\varepsilon_2 \in \{\pm 1\} \\ \varepsilon_1\varepsilon_2 = 1}} \sum_{(u,v) \in Z^2 \cap R^{\varepsilon_1,\varepsilon_2}(\frac{B}{kt})} r(t^2 L^+ C^+),$$

where we have written $L^+ = \varepsilon_1 L$ and $C^+ = \varepsilon_2 C$.

In what follows it will be convenient to write

$$\omega(a_1, \dots, a_k) = \omega(\gcd(a_1, \dots, a_k)),$$

where $\omega(n) = \sum_{p \mid n} 1$. We would now like to break the summand into a part involving $t^2$, a part involving $L^+$ and a part involving $C^+$. For this we call upon the following result, which is established along precisely the same lines as [3, Lemme 10], where the analogous formula for the divisor function is established.

LEMMA 9: *Let $n_1, n_2, n_3 \in \mathbb{N}$. Then we have*

$$r(n_1n_2n_3) = \sum_{d_id_j|n_k} \frac{\chi(d_1d_2d_3)\mu(d_1)\mu(d_2d_3)}{2^{\omega(d_2d_3,n_2,n_3)+4}} r\Big(\frac{n_1}{d_2d_3}\Big) r\Big(\frac{n_2}{d_1d_3}\Big) r\Big(\frac{n_3}{d_1d_2}\Big),$$

*where the indices $\{i,j,k\}$ run over permutations of the set $\{1,2,3\}$.*

Applying Lemma 9, we conclude that

$$r(t^2L^+C^+) = \sum_{d_1d_2|t^2} \sum_{\substack{d_1d_3|L\\ d_2d_3|C}} \frac{\chi(d_1d_2d_3)\mu(d_3)\mu(d_1d_2)}{2^{\omega(d_1d_2,L,C)+4}} r\Big(\frac{t^2}{d_1d_2}\Big) r\Big(\frac{L^+}{d_1d_3}\Big) r\Big(\frac{C^+}{d_2d_3}\Big).$$

Write $d = d_1d_2$ and note that $d \mid t$ for any value of $d$ producing a non-zero summand. In particular we will only be interested in values of $d \in \mathfrak{D}$, so that $\chi(d) = 1$. Writing $t = ds$, we deduce that

$$N(B) = \frac{1}{2^5} \sum_{\substack{dk \leqslant B\\ d,k\in\mathfrak{D}}} \mu(d)\mu(k) \sum_{\substack{s \leqslant \frac{B}{dk}\\ s\in\mathfrak{D}}} r(ds^2) \sum_{\substack{\mathbf{d}\in\mathbb{N}^3\\ d=d_1d_2}} \chi(d_3)\mu(d_3)\mathcal{S}_{\mathbf{d}}\Big(\frac{B}{dsk}\Big),$$

where

$$\mathcal{S}_{\mathbf{d}}(T) = \sum_{\substack{\varepsilon_1,\varepsilon_2\in\{\pm1\}\\ \varepsilon_1\varepsilon_2=1}} \sum_{\substack{(u,v)\in Z^2\cap R^{\varepsilon_1,\varepsilon_2}(T)\\ d_1d_3|L,\ d_2d_3|C}} \frac{r(\frac{L^+}{d_1d_3}) r(\frac{C^+}{d_2d_3})}{2^{\omega(d,L,C)}},$$

for any $T \geqslant 1$. Now the inner sum vanishes unless $d_3 \mid \gcd(L(u,v), C(u,v))$, with $(u,v)$ a primitive integer vector. In particular it follows that $d_3 \mid \Delta$, the resultant of $L$ and $C$, whence $d_3 = O(1)$.

For given $d \in \mathbb{N}$ we let

$$f_d(n) = \sum_{n=ab} \mu(a) r(db^2). \tag{5.4}$$

We may now write

$$N(B) = \frac{1}{2^5} \sum_{\substack{dn \leqslant B\\ d,n\in\mathfrak{D}}} \mu(d) f_d(n) \sum_{\substack{\mathbf{d}\in\mathbb{N}^3\\ d=d_1d_2\\ d_3|\Delta}} \chi(d_3)\mu(d_3)\mathcal{S}_{\mathbf{d}}\Big(\frac{B}{dn}\Big).$$

Recycling the observation that any common divisor of $L(u,v)$ and $C(u,v)$ must divide $\Delta$, we obtain

$$\begin{aligned}\mathcal{S}_{\mathbf{d}}(T) &= \sum_{\substack{\varepsilon_1,\varepsilon_2\in\{\pm1\}\\ \varepsilon_1\varepsilon_2=1}} \sum_{k\mid\gcd(\Delta,d)} \frac{1}{2^{\omega(k)}} \sum_{\substack{(u,v)\in Z^2\cap R^{\varepsilon_1,\varepsilon_2}(T)\\ d_1d_3\mid L,\ d_2d_3\mid C\\ k=\gcd(d,L,C)}} r\Big(\frac{L^+}{d_1d_3}\Big)r\Big(\frac{C^+}{d_2d_3}\Big)\\ &= \sum_{\substack{\varepsilon_1,\varepsilon_2\in\{\pm1\}\\ \varepsilon_1\varepsilon_2=1}} \sum_{kk'\mid\gcd(\Delta,d)} \frac{\mu(k')}{2^{\omega(k)}} \sum_{\substack{(u,v)\in Z^2\cap R^{\varepsilon_1,\varepsilon_2}(T)\\ [d_1d_3,kk']\mid L\\ [d_2d_3,kk']\mid C}} r\Big(\frac{L^+}{d_1d_3}\Big)r\Big(\frac{C^+}{d_2d_3}\Big).\end{aligned}$$

Finally, we wish to remove the coprimality condition on $(u,v)$ using the Möbius function. Let us define

$$L_\ell = \ell L^+ = \ell\varepsilon_1 L, \quad C_\ell = \ell^3 C^+ = \ell^3\varepsilon_2 C \tag{5.5}$$

for any $\ell\in\mathbb{N}$. It follows that the inner sum over $(u,v)$ is equal to $\sum_{\ell\leqslant\sqrt{T}}\mu(\ell)\mathcal{U}(\ell^{-2}T)$, where if $\mathbf{k}=(k,k')$ then

$$\mathcal{U}(T)=\mathcal{U}^{\varepsilon_1,\varepsilon_2}_{\mathbf{d},\mathbf{k},\ell}(T)= \sum_{\substack{(x,y)\in\mathbb{Z}^2\cap R^{\varepsilon_1,\varepsilon_2}(T)\\ [d_1d_3,kk']\mid L_\ell\\ [d_2d_3,kk']\mid C_\ell}} r\Big(\frac{L_\ell(x,y)}{d_1d_3}\Big)r\Big(\frac{C_\ell(x,y)}{d_2d_3}\Big). \tag{5.6}$$

We may summarise our investigation as follows.

LEMMA 10: *There exists an absolute constant $c>0$ such that*

$$\begin{aligned}N(B) =& \frac{1}{2^5}\sum_{\ell=1}^{\infty}\mu(\ell)\sum_{d\in\mathfrak{D}}\mu(d)\sum_{\substack{n\leqslant N\\ n\in\mathfrak{D}}} f_d(n) \sum_{\substack{\varepsilon_1,\varepsilon_2\in\{\pm1\}\\ \varepsilon_1\varepsilon_2=1}} \sum_{kk'\mid\gcd(\Delta,d)} \frac{\mu(k')}{2^{\omega(k)}}\\ &\times \sum_{\substack{\mathbf{d}\in\mathbb{N}^3\\ d=d_1d_2\\ d_3\mid\Delta}} \chi(d_3)\mu(d_3)\mathcal{U}\Big(\frac{B}{d\ell^2n}\Big),\end{aligned}$$

*where $N=\frac{cB}{d^{\frac54}\ell}$ and $\mathcal{U}(T)=\mathcal{U}^{\varepsilon_1,\varepsilon_2}_{\mathbf{d},\mathbf{k},\ell}(T)$ is given by* (5.6).

*Proof.* In view of our preceding manipulations, the statement of the lemma is obviously true with $N=B/d\ell^2$ in the summation over $n$. To see that we may take $N=cB/d^{\frac54}\ell$ for some absolute constant $c>0$, we observe that $\mathcal{U}(T)=0$ unless $d_1\ll\ell^3T^{\frac32}$ and $d_2\ll\ell T^{\frac12}$. Taking $T=B/d\ell^2n$, it follows that

$d = d_1 d_2 \ll B^2/d^2n^2$, whence $d^{\frac{3}{2}}n \ll B$. But we also have $d\ell^2 n \leqslant B$, whence in fact $d^{\frac{5}{4}}\ell n \ll B$, as required. ∎

The groundwork is now laid for an investigation of $\mathcal{U}(T)$ for appropriate values of the parameters. In effect, the thrust of this section has been concerned with passing from solutions of a single equation $y^2 + z^2 = t^2 L(u,v)C(u,v)$, to solutions of

$$\ell L(u,v) = \delta_1(y_1^2 + z_1^2), \quad \ell^3 C(u,v) = \delta_2(y_2^2 + z_2^2),$$

for varying $\delta_1, \delta_2 \in \mathbb{Z}$. This corresponds to a simple descent process and the pair of equations defines an intermediate torsor above the Châtelet surface $X$.

## 6. Analysis of $\mathcal{U}(T)$

In this section we will study $\mathcal{U}(T) = \mathcal{U}^{\varepsilon_1,\varepsilon_2}_{\mathbf{d},\mathbf{k},\ell}(T)$, as given by (5.6). We will work with the sets

$$\Lambda(\mathbf{D}) = \Lambda(\mathbf{D}; L, C) = \{\mathbf{x} \in \mathbb{Z}^2 : D_1 \mid L(\mathbf{x}),\ D_2 \mid C(\mathbf{x})\},$$
$$\Lambda^*(\mathbf{D}) = \Lambda^*(\mathbf{D}; L, C) = \{\mathbf{x} \in \Lambda(\mathbf{D}; L, C) : \gcd(D_1 D_2, \mathbf{x}) = 1\},$$

for any $\mathbf{D} \in \mathbb{N}^2$. Let us write

$$e_1 = d_1 d_3, \quad e_2 = d_2 d_3, \quad E_1 = [d_1 d_3, kk'], \quad E_2 = [d_2 d_3, kk']. \tag{6.1}$$

Clearly $e_i, E_i$ are all odd and $e_i \mid E_i$. Let $\mathcal{R} = R^{\varepsilon_1,\varepsilon_2}(1)$, so that $\sqrt{T}\mathcal{R} = R^{\varepsilon_1,\varepsilon_2}(T)$. We may therefore write

$$\mathcal{U}(T) = \sum_{\mathbf{x} \in \Lambda(\mathbf{E}; L_\ell, C_\ell) \cap \sqrt{T}\mathcal{R}} r\Big(\frac{L_\ell(\mathbf{x})}{e_1}\Big) r\Big(\frac{C_\ell(\mathbf{x})}{e_2}\Big),$$

where $L_\ell, C_\ell$ are given by (5.5). Ultimately we wish to apply Theorem 2 to estimate this sum. However, the latter result involves a sum over points of $\mathbb{Z}^2$ rather than points of $\Lambda(\mathbf{E}; L_\ell, C_\ell)$. We will circumvent this difficulty with a change of variables.

The first task is to restrict attention to the case in which each $E_1$ (resp. $E_2$) is coprime to the coefficients of $L_\ell$ (resp. $C_\ell$). We let $\ell_1, \ell_2 \in \mathbb{N}$ and $L^*, C^*$ be primitive forms such that $L_\ell = \ell_1 L^*$ and $C_\ell = \ell_2 C^*$. In particular $\ell \mid \ell_1, \ell^3 \mid \ell_2$ and $\ell^{-1}\ell_1, \ell^{-3}\ell_2 \ll 1$. Then $\Lambda(\mathbf{E}; L_\ell, C_\ell) = \Lambda(\mathbf{E}'; L^*, C^*)$, with

$$E_1' = \frac{E_1}{\gcd(E_1, \ell_1)}, \quad E_2' = \frac{E_2}{\gcd(E_2, \ell_2)}. \tag{6.2}$$

Define the function $\psi : \mathbb{N}^2 \to \mathbb{N}$ multiplicatively via

$$\psi(p^{\alpha_1}, p^{\alpha_2}) = p^{\max\{\alpha_1, \lceil \frac{\alpha_2}{3} \rceil\}}.$$

Breaking the set according to the value of $\gcd(\psi(\mathbf{E}'), \mathbf{x})$ one deduces that

$$\Lambda(\mathbf{E}'; L^*, C^*) = \bigsqcup_{h \mid \psi(\mathbf{E}')} h\Lambda^*(\mathbf{E}''; L^*, C^*) = \bigsqcup_{h \mid \psi(\mathbf{E}')} h\Lambda^*(\mathbf{E}''),$$

where

$$E_1'' = \frac{E_1'}{\gcd(E_1', h)}, \quad E_2'' = \frac{E_2'}{\gcd(E_2', h^3)}. \tag{6.3}$$

Here one notes that $h^{-1}\psi(\mathbf{E}') = \psi(\mathbf{E}'')$ and furthermore $\gcd(\psi(\mathbf{E}''), \mathbf{x}) = 1$ if and only if $\gcd(E_1'' E_2'', \mathbf{x}) = 1$. Replacing $\mathbf{x}$ by $h\mathbf{x}$ we note that

$$r\Big(\frac{h\ell_1 L^*(\mathbf{x})}{e_1}\Big) r\Big(\frac{h^3 \ell_2 C^*(\mathbf{x})}{e_2}\Big) = r\Big(\frac{\ell_1' L^*(\mathbf{x})}{e_1'}\Big) r\Big(\frac{\ell_2' C^*(\mathbf{x})}{e_2'}\Big),$$

where

$$\ell_1' = \frac{\ell_1 h}{\gcd(e_1, h)}, \quad \ell_2' = \frac{\ell_2 h^3}{\gcd(e_2, h^3)},$$

and

$$e_1' = \frac{e_1}{\gcd(e_1, h)}, \quad e_2' = \frac{e_2}{\gcd(e_2, h^3)}.$$

It now follows that

$$\mathcal{U}(T) = \sum_{h \mid \psi(\mathbf{E}')} \sum_{\mathbf{x} \in \Lambda^*(\mathbf{E}'') \cap h^{-1}\sqrt{T}\mathcal{R}} r\Big(\frac{\ell_1' L^*(\mathbf{x})}{e_1'}\Big) r\Big(\frac{\ell_2' C^*(\mathbf{x})}{e_2'}\Big).$$

We let $e' = e_1' e_2'$, $E' = E_1' E_2'$ and $E'' = E_1'' E_2''$.

In $\Lambda^*(\mathbf{E}'')$ we define an equivalence relation $\mathbf{x} \sim \mathbf{y}$ if and only if there exists $\lambda \in \mathbb{Z}$ such that

$$\mathbf{x} \equiv \lambda \mathbf{y} \,(\mathrm{mod}\, E'').$$

Note that any such $\lambda$ must be coprime to $E''$. This relation allows us to partition $\Lambda^*(\mathbf{E}'')$ into disjoint equivalence classes. We denote by $\mathcal{U}(\mathbf{E}'')$ the set of these equivalence classes. We claim that

$$\#\mathcal{U}(\mathbf{D}) \ll (D_1 D_2 D_3)^{\varepsilon} \tag{6.4}$$

for any $\mathbf{D} \in \mathbb{N}^2$. To see this we note that

$$\#\mathcal{U}(\mathbf{D}) = \frac{\varrho^*(\mathbf{D})}{\varphi(D_1 D_2)} = \prod_{p^{\nu_i} \| D_i} \frac{\varrho^*(p^{\nu_1}, p^{\nu_2})}{\varphi(p^{\nu_1 + \nu_2})},$$

where $\varrho^*(\mathbf{D}) = \varrho^*(\mathbf{D}; L^*, C^*)$ is given multiplicatively as in (2.6). Applying (2.10) we easily deduce (6.4).

When $\mathbf{y} \in \mathcal{A}$ for $\mathcal{A} \in \mathcal{U}(\mathbf{E}'')$, we have

$$\mathcal{A} = \{\mathbf{x} \in \mathbb{Z}^2 : \mathbf{x} \equiv \lambda \mathbf{y} \,(\mathrm{mod}\, E'') \text{ with } \lambda \in \mathbb{Z} \text{ and } \gcd(\lambda, E'') = 1\}.$$

When $\mathcal{A} \in \mathcal{U}(\mathbf{E}'')$ and $\mathbf{y}_0 \in \mathcal{A}$, we set

$$G(\mathcal{A}) = \{\mathbf{x} \in \mathbb{Z}^2 : \exists \lambda \in \mathbb{Z} \text{ such that } \mathbf{x} \equiv \lambda \mathbf{y}_0 \,(\mathrm{mod}\, E'')\}.$$

This defines a sublattice of $\mathbb{Z}^2$ of rank 2 and determinant $E''$. Moreover, the definition is independent of $\mathbf{y}_0$. We conclude that

$$\mathcal{U}(T) = \sum_{h \mid \psi(\mathbf{E}')} \sum_{\mathcal{A} \in \mathcal{U}(\mathbf{E}'')} \sum_{e \mid E''} \mu(e) S(T, \mathcal{A}, e) \tag{6.5}$$

where

$$S(T, \mathcal{A}, e) = \sum_{\mathbf{x} \in G_e(\mathcal{A}) \cap h^{-1}\sqrt{T}\mathcal{R}} r\Big(\frac{\ell_1' L^*(\mathbf{x})}{e_1'}\Big) r\Big(\frac{\ell_2' C^*(\mathbf{x})}{e_2'}\Big),$$

with

$$G_e(\mathcal{A}) = G(\mathcal{A}) \cap \{\mathbf{x} \in \mathbb{Z}^2 : e \mid \mathbf{x}\} = \{\mathbf{x} \in \mathbb{Z}^2 : \exists a \in e\mathbb{Z} \text{ s.t. } \mathbf{x} \equiv a\mathbf{y}_0 \,(\mathrm{mod}\, E'')\}.$$

We have therefore arrived at summation conditions running over a lattice $G_e(\mathcal{A})$ of determinant $eE''$. We claim that

$$\det G_e(\mathcal{A}) \gg \frac{de}{\gcd(d, h\ell)}. \tag{6.6}$$

For this we note from (6.1), (6.2) and (6.3) that

$$E'' = E_1'' E_2'' \geqslant \frac{[d_1 d_3, kk']}{\gcd([d_1 d_3, kk'], h\ell_1)} \cdot \frac{[d_2 d_3, kk']}{\gcd([d_2 d_3, kk'], h^3 \ell_2)}.$$

Now we have seen in Lemma 10 that $d = d_1 d_2$ is square-free and $d_3, k, k' \ll 1$. Since $\ell_1 \ll \ell$ and $\ell_2 \ll \ell^3$ it easily follows that

$$E'' \gg \frac{d_1 d_2}{\gcd(d_1, h\ell_1) \gcd(d_2, h^3 \ell_2)} \gg \frac{d}{\gcd(d, h\ell)},$$

as required for (6.6).

We are now led to make a change of variables $\mathbf{x} = \mathbf{M}\mathbf{v}$ for any $\mathbf{x} \in G_e(\mathcal{A})$, where $\mathbf{M} = (\mathbf{m}_1, \mathbf{m}_2)$ is the matrix formed from a minimal basis for the lattice. In particular, if $s_1 \leqslant s_2$ are the successive minima of $G_e(\mathcal{A})$ with respect to the norm $|\cdot|$, then $s_i = |\mathbf{m}_i|$ for $i = 1, 2$ and $s_1 s_2$ has order of magnitude $eE''$. Moreover, according to Davenport's work in the geometry of numbers [9, Lemma 5], we will have $v_i \ll s_i^{-1}|\mathbf{x}|$ whenever $\mathbf{x} \in G_e(\mathcal{A})$ is written as

$\mathbf{x} = v_1\mathbf{m}_1 + v_2\mathbf{m}_2$. On defining the region $\mathcal{R}_{\mathbf{M}} = \{\mathbf{v} \in \mathbb{R}^2 : \mathbf{M}\mathbf{v} \in h^{-1}\mathcal{R}\}$, we observe that

$$\mathrm{vol}(\mathcal{R}_{\mathbf{M}}) = \frac{\mathrm{vol}(\mathcal{R})}{h^2|\det \mathbf{M}|} = \frac{\mathrm{vol}(\mathcal{R})}{h^2 eE''}. \tag{6.7}$$

We may now write

$$S(T, \mathcal{A}, e) = \sum_{\mathbf{v} \in \mathbb{Z}^2 \cap \sqrt{T}\mathcal{R}_{\mathbf{M}}} r(M_1(\mathbf{v}))r(M_2(\mathbf{v})) \tag{6.8}$$

with

$$M_1(\mathbf{v}) = \frac{\ell_1' L^*(\mathbf{M}\mathbf{v})}{e_1'}, \quad M_2(\mathbf{v}) = \frac{\ell_2' C^*(\mathbf{M}\mathbf{v})}{e_2'}.$$

Our analysis of $S(T, \mathcal{A}, e)$ will now involve two aspects: a uniform upper bound and an asymptotic formula. In the first instance, therefore, we require an upper bound for this sum which is uniform in $d = d_1 d_2$ and $\ell$. Our principal tool will be previous work of the authors [1], which is concerned with the average order of arithmetic functions ranging over the values taken by binary forms. As usual we will allow all of our implied constants to depend upon the coefficients of the forms $L$ and $C$. In particular we have $d_3 \ll 1$. We will establish the following result.

LEMMA 11: *Let $\varepsilon > 0$ and let $d$ be square-free. Then we have*

$$\mathcal{U}(T) \ll (d\ell)^{\varepsilon} \gcd(d, \ell)\Big(\frac{T}{d} + T^{\frac{1}{2}+\varepsilon}\Big).$$

*Proof.* Let $r_2(n)$ be defined multiplicatively via

$$r_2(p^j) = \begin{cases} 1 + \chi(p), & \text{if } j = 1 \text{ and } p \nmid 6dd_3\Delta\Delta'\ell, \\ (1+j)^2, & \text{otherwise,} \end{cases}$$

where $\Delta, \Delta'$ are as in (2.5). It follows from (6.8) that

$$S(T, \mathcal{A}, e) \leqslant 2^4 \sum_{\substack{\mathbf{v} \in \mathbb{Z}^2 \\ v_1 \ll V_1,\ v_2 \ll V_2}} r_2\big(M_1(\mathbf{v})M_2(\mathbf{v})\big),$$

where $V_i = (hs_i)^{-1}\sqrt{T}$ for $i = 1, 2$.

It is obvious that $r_2$ belongs to the class of non-negative arithmetic functions considered in [1]. An application of [1, Corollary 1] therefore reveals that

$$S(T, \mathcal{A}, e) \ll (d\ell)^{\varepsilon}(V_1V_2E + V_1^{1+\varepsilon}) \ll (d\ell)^{\varepsilon}\Big(\frac{T}{h^2 s_1 s_2}E + \frac{T^{\frac{1}{2}+\varepsilon}}{hs_1}\Big),$$

for any $\varepsilon > 0$, where

$$E = \prod_{p \leqslant V_2} \Big(1 + \frac{\varrho_{M_2(x,1)}(p)\chi(p)}{p}\Big).$$

It follows from Lemma 2 that $E \leqslant A^{\omega(d\ell)} \ll (d\ell)^{\varepsilon}$ for an appropriate constant $A \geqslant 1$. Recalling that $s_1 s_2 \gg eE''$, we therefore conclude from (6.6) that

$$S(T, \mathcal{A}, e) \ll (d\ell)^{\varepsilon}\Big(\frac{T\gcd(d, h\ell)}{deh^2} + \frac{T^{\frac{1}{2}+\varepsilon}}{h}\Big).$$

Inserting this into (6.5) now yields

$$\begin{aligned}\mathcal{U}(T) &\ll (d\ell)^{\varepsilon} \sum_{h \mid \psi(\mathbf{E}')} \frac{\gcd(d, h)}{h} \#\mathcal{U}(\mathbf{E}'')\Big(\frac{T\gcd(d, \ell)}{d} + T^{\frac{1}{2}+\varepsilon}\Big)\\ &\ll (d\ell)^{\varepsilon}\gcd(d, \ell)\Big(\frac{T}{d} + T^{\frac{1}{2}+\varepsilon}\Big),\end{aligned}$$

by (6.4). This completes the proof of Lemma 11. ∎

We now turn to an asymptotic formula for $\mathcal{U}(T) = \mathcal{U}^{\varepsilon_1,\varepsilon_2}_{\mathbf{d},\mathbf{k},\ell}(T)$, as given by (6.5) and (6.8). Whereas in the previous lemma we sought uniformity in $d = d_1 d_2$ and $\ell$, we will now allow all of our implied constants to depend in any way upon $d, \ell$ and the coefficients of $L$ and $C$. It is clear that $\mathcal{R}_{\mathbf{M}}$ and $M_1, M_2$ satisfy the necessary conditions for an application of Theorem 2. Put

$$K_p(\mathbf{M}) = \Big(1 - \frac{\chi(p)}{p}\Big)^2 \sum_{\nu_1,\nu_2 \geqslant 0} \frac{\chi(p^{\nu_1+\nu_2})\varrho(p^{\nu_1}, p^{\nu_2}; M_1, M_2)}{p^{2\nu_1+2\nu_2}}$$

for $p > 2$ and

$$K_2(\mathbf{M}) = 4 \lim_{n\to\infty} 2^{-2n} \# \left\{ \mathbf{x} \in (\mathbb{Z}/2^n\mathbb{Z})^2 : \begin{array}{l} M_1(\mathbf{x}) \in \mathcal{E} \pmod{2^n} \\ M_2(\mathbf{x}) \in \mathcal{E} \pmod{2^n} \end{array} \right\}.$$

Then once combined with (6.5) and (6.7), Theorem 2 leads to the following result.

LEMMA 12: *Let $\varepsilon > 0$. Then we have*

$$\mathcal{U}(T) = \pi^2 W^{\varepsilon_1,\varepsilon_2}(\mathbf{d}, \mathbf{k}, \ell)\operatorname{vol}(R^{\varepsilon_1,\varepsilon_2}(1))T + O\big(T(\log T)^{-\eta+\varepsilon}\big),$$

*where the implied constant depends on $d, \ell, L, C$, and*

$$W^{\varepsilon_1,\varepsilon_2}(\mathbf{d}, \mathbf{k}, \ell) = \sum_{h \mid \psi(\mathbf{E}')} \sum_{\mathcal{A} \in \mathcal{U}(\mathbf{E}'')} \sum_{e \mid E''} \frac{\mu(e)}{h^2 eE''} \prod_p K_p(\mathbf{M}).$$

It will be useful to have an expression for $W(\mathbf{d},\ell)$ as an Euler product. Following the argument in [3, §6] almost verbatim one is led to the conclusion that

$$W^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k},\ell)=\prod_p W_p^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k},\ell),$$

where for $p>2$,

$$W_p^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k},\ell)=\Big(1-\frac{\chi(p)}{p}\Big)^2\sum_{\nu_1,\nu_2\geqslant 0}\frac{\chi(p^{\nu_1+\nu_2})\varrho(p^{N_1},p^{N_2};L_\ell,C_\ell)}{p^{2N_1+2N_2}}, \tag{6.9}$$

with $N_i=\max\{v_p(E_i),\nu_i+v_p(e_i)\}$ for $i=1,2$, and

$$W_2^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k},\ell)=4\lim_{n\to\infty}2^{-2n}\#\left\{\mathbf{x}\in(\mathbb{Z}/2^n\mathbb{Z})^2:\ \begin{array}{l}\ell L(\mathbf{x})\in\varepsilon_1 d_3\mathcal{E}\,(\mathrm{mod}\,2^n)\\ \ell^3 C(\mathbf{x})\in\varepsilon_1 d_3\mathcal{E}\,(\mathrm{mod}\,2^n)\end{array}\right\}. \tag{6.10}$$

We have used here the fact that $d_1\equiv d_2\equiv 1\,(\mathrm{mod}\,4)$ and $\varepsilon_1\varepsilon_2=1$. In our work we will also need a good upper bound for the constant $W^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k},\ell)$ which is uniform in $d$ and $\ell$. This is recorded in the following result.

LEMMA 13: *We have* $W^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k},\ell)\ll d^{-\frac{1}{6}+\varepsilon}\ell^{\varepsilon}$ *for any* $\varepsilon>0$.

*Proof.* Building on the above Euler product representation of $W^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k},\ell)$, it is clear that $|W_2^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k},\ell)|\leqslant 4$. Thus we focus our attention on the factors corresponding to odd primes. When $p>2$, part (3) of Lemma 3 implies that

$$|W_p^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k},\ell)|\ll\sum_{\nu_1,\nu_2\geqslant 0}\frac{\min\{p^{N_1+2N_2},p^{2N_1+\frac{5N_2}{3}}\}}{p^{2N_1+2N_2}}\ll\sum_{\nu_1,\nu_2\geqslant 0}\frac{1}{p^{\frac{N_1}{2}+\frac{N_2}{6}}}.$$

Suppose that $v_p(d_1)=\delta_1$ and $v_p(d_2)=\delta_2$. Since $d=d_1d_2$ is square-free we may assume that $\delta_1+\delta_2=1$ if $p\mid d$. Moreover, $N_1\geqslant\delta_1+\nu_1$ and $N_2\geqslant\delta_2+\nu_2$. We conclude that

$$\prod_{p\mid d}|W_p^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k},\ell)|\ll d^{\varepsilon}\prod_{p\mid d}p^{-\frac{\delta_1+\delta_2}{6}}\sum_{\nu_1,\nu_2\geqslant 0}p^{-\frac{\nu_1}{2}-\frac{\nu_2}{6}}\ll d^{-\frac{1}{6}+\varepsilon}.$$

Taking $N_i\geqslant\nu_i$ it also follows that $\prod_{p\mid D}|W_p^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k},\ell)|\ll D^{\varepsilon}$, for any odd $D\in\mathbb{N}$. Finally, the analysis in the proof of Lemma 8, which is based on repeated applications of Lemma 3, furnishes the bound

$$\prod_{p\nmid 2d\ell\Delta\Delta' c_0}|W_p(\mathbf{d},\mathbf{k},\ell)|\ll(d\ell)^{\varepsilon}.$$

Putting everything together therefore concludes the proof of the lemma. ∎

## 7. Concluding steps

We are now ready to draw to a close our proof of Theorem 1, for which we begin with some technical estimates. Recall the definition (5.3) of the set $\mathfrak{D}$ and the definition (5.4) of the function $f_d(n)$. We will need the following easy result.

LEMMA 14: *Let $d \in \mathfrak{D}$ be square-free. Then we have*

$$\sum_{\substack{n \leqslant x \\ n \in \mathfrak{D}}} \frac{f_d(n)}{n} = \frac{r(d)\varphi^\dagger(d)}{\pi}\Big(\log x + O\big(\log^3(2+\omega(d))\big)\Big),$$

*where $\varphi^\dagger(d) = \prod_{p \mid d}(1+\frac{1}{p})^{-1}$.*

*Proof.* The proof of Lemma 14 involves a straightforward consideration of the corresponding Dirichlet series $F_d(s) = \sum_{n \in \mathfrak{D}} f_d(n) n^{-s}$. Let $r_0(n) = \frac{1}{4} r(n)$. It is easy to see that

$$F_d(s) = 4 \sum_{m \in \mathfrak{D}} \frac{\mu(m)}{m^s} \sum_{n \in \mathfrak{D}} \frac{r_0(dn^2)}{n^s},$$

Let $\delta = \delta_p = v_p(d)$. Then for square-free $d \in \mathfrak{D}$ we have $\delta \in \{0,1\}$ and $\delta = 1$ if and only if $p \mid d$ and $p \equiv 1 \pmod 4$. We now have

$$\begin{aligned} F_d(s) &= 4 \prod_{p \equiv 1 \,(\mathrm{mod}\, 4)} \Big(1 - \frac{1}{p^s}\Big) \prod_{p \equiv 1 \,(\mathrm{mod}\, 4)} \sum_{\nu \geqslant 0} \frac{1+\delta+2\nu}{p^{\nu s}} \\ &= 4 \prod_{p \equiv 1 \,(\mathrm{mod}\, 4)} \Big(\frac{1+p^{-s}}{1-p^{-s}}\Big) \prod_{p \equiv 1 \,(\mathrm{mod}\, 4)} \Big(\frac{1+\delta+(1-\delta)p^{-s}}{1+p^{-s}}\Big) \\ &= \frac{4\zeta(s)L(s,\chi)}{(1+2^{-s})\zeta(2s)} H_d(s), \end{aligned}$$

where

$$H_d(s) = \prod_{p \mid d} \Big(\frac{2}{1+p^{-s}}\Big) = r_0(d) \prod_{p \mid d} \Big(1 + \frac{1}{p^s}\Big)^{-1}.$$

Noting that $H_1(s) = 1$ we clearly have $F_d(s) = F_1(s) H_d(s)$.

The Dirichlet series $F_1(s)$ is meromorphic in the region $\Re e(s) > \frac{1}{2}$, with a simple pole at $s = 1$. Moreover, there is an arithmetic function $h_d(n)$, arising from the Dirichlet series $H_d(s)$, such that $f_d = f_1 * h_d$. On applying a Tauberian theorem one easily deduces that the statement of Lemma 14 is true when $d = 1$.

To see the general case we note that

$$\sum_{n\leqslant x}\frac{f_d(n)}{n}=\sum_{m\leqslant x}\frac{h_d(m)}{m}\sum_{n\leqslant \frac{x}{m}}\frac{f_1(n)}{n}=\sum_{m\leqslant x}\frac{h_d(m)}{m}\Big(\frac{4\log x}{\pi}+O(\log 2m)\Big).$$

Here

$$\begin{aligned}\sum_{m=1}^{\infty}\frac{|h_d(m)|\log 2m}{m}&\leqslant r_0(d)\varphi^\dagger(d)^{-1}\Big(1+\sum_{p\mid d}\frac{\log p}{p}\Big)\\&\ll r(d)\varphi^\dagger(d)\varphi^\dagger(d)^{-2}\log(2+\omega(d))\\&\ll r(d)\varphi^\dagger(d)\log^3(2+\omega(d)),\end{aligned}$$

since

$$\sum_{p\mid d}\frac{\log p}{p}\leqslant\sum_{j\leqslant\omega(d)}\frac{\log p_j}{p_j}\ll\log(2+\omega(d)).$$

On inserting this into the previous formula, we therefore complete the proof of the lemma since $H_d(1)=r_0(d)\varphi^\dagger(d)$. ∎

Building on Lemma 14, we may record the inequalities

$$\sum_{\substack{n\leqslant x\\ n\in\mathfrak{D}}}\frac{|f_d(n)|}{n^\theta}\leqslant x^{1-\theta}\sum_{\substack{n\leqslant x\\ n\in\mathfrak{D}}}\frac{|f_d(n)|}{n}\ll d^\varepsilon x^{1-\theta}\log x, \tag{7.1}$$

for any $\varepsilon>0$ and $0<\theta\leqslant 1$. For the deduction of Theorem 1, we wish to incorporate the asymptotic formula in Lemma 12 into our expression for $N(B)$ in Lemma 10. Note that there is no uniformity in any of the parameters $\mathbf{d},\mathbf{k},\ell$ that feature in Lemma 12. Let us set

$$S(B)=S^{\varepsilon_1,\varepsilon_2}_{\mathbf{d},\mathbf{k},\ell}(B)=\sum_{\substack{n\leqslant N\\ n\in\mathfrak{D}}}f_d(n)\mathcal{U}\Big(\frac{B}{d\ell^2 n}\Big),$$

with $N=cB/d^{\frac{5}{4}}\ell$ for some absolute constant $c>0$, so that

$$N(B)=\frac{1}{2^5}\sum_{\ell=1}^{\infty}\mu(\ell)\sum_{d\in\mathfrak{D}}\mu(d)\sum_{\substack{\varepsilon_1,\varepsilon_2\in\{\pm1\}\\ \varepsilon_1\varepsilon_2=1}}\sum_{\substack{\mathbf{d}\in\mathbb{N}^3\\ d=d_1d_2\\ d_3\mid\Delta}}\chi(d_3)\mu(d_3)\sum_{kk'\mid\gcd(\Delta,d)}\frac{\mu(k')}{2^{\omega(k)}}S(B).$$

Let

$$\begin{aligned}&E^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k},\ell)\\&\quad=\frac{1}{B\log B}\Big|S(B)-\frac{\pi W^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k},\ell)\operatorname{vol}(R^{\varepsilon_1,\varepsilon_2}(1))r(d)\varphi^\dagger(d)B\log B}{d\ell^2}\Big|.\end{aligned}$$

Then it follows from Lemmas 12 and 14 that for fixed $\mathbf{d}, \mathbf{k}, \ell$ we have

$$E^{\varepsilon_1,\varepsilon_2}(\mathbf{d}, \mathbf{k}, \ell) \to 0$$

as $B \to \infty$. On the other hand, we conclude from (7.1) and Lemmas 11 and 13 that

$$E^{\varepsilon_1,\varepsilon_2}(\mathbf{d}, \mathbf{k}, \ell) \ll (d\ell)^{\varepsilon} \gcd(d, \ell)\Big(\frac{1}{d^2\ell^2} + \frac{1}{d^{\frac{9}{8}}\ell^{\frac{3}{2}}} + \frac{1}{d^{\frac{7}{6}}\ell^2}\Big) \ll (d\ell)^{\varepsilon} \frac{\gcd(d, \ell)}{d^{\frac{9}{8}}\ell^{\frac{3}{2}}},$$

uniformly in $d, \ell$ and $B$. Note that

$$\sum_{\ell} \sum_{d} \sum_{\varepsilon_1,\varepsilon_2} \sum_{\mathbf{d}} \sum_{\mathbf{k}} E^{\varepsilon_1,\varepsilon_2}(\mathbf{d}, \mathbf{k}, \ell) \ll 1.$$

Writing $r_0(n) = \frac{1}{4}r(n)$, it therefore follows from the dominated convergence of this sum that as $B \to \infty$ we have $N(B) \sim c_0 B \log B$, with

$$\begin{aligned} c_0 = &\frac{\pi}{2^3} \sum_{\ell=1}^{\infty} \frac{\mu(\ell)}{\ell^2} \sum_{d \in \mathfrak{D}} \frac{\mu(d) r_0(d) \varphi^{\dagger}(d)}{d} \sum_{\substack{\varepsilon_1,\varepsilon_2 \in \{\pm 1\} \\ \varepsilon_1\varepsilon_2 = 1}} \operatorname{vol}(R^{\varepsilon_1,\varepsilon_2}(1)) \\ &\times \sum_{\substack{\mathbf{d} \in \mathbb{N}^3 \\ d = d_1 d_2 \\ d_3 \mid \Delta}} \chi(d_3)\mu(d_3) \sum_{kk' \mid \gcd(\Delta, d)} \frac{\mu(k')}{2^{\omega(k)}} W^{\varepsilon_1,\varepsilon_2}(\mathbf{d}, \mathbf{k}, \ell). \end{aligned} \tag{7.2}$$

Now let $c_X$ be the constant predicted by Peyre [17]. In order to complete the proof of Theorem 1 it remains to show that $c_0 = c_X$. Given the general strategy in our earlier work [4], we will be brief. Relating the value of the constant $c_X$ to the count on the torsor $\mathcal{T}$ considered in (5.2), one finds that

$$c_X = \omega_\infty \prod_p \omega_p,$$

where $\omega_\infty$ and $\omega_p$ denote the local densities associated to $\mathcal{T}$ taken with respect to the Leray measure. Using symmetry to restrict to the quadrant in which $y > 0$ and $z > 0$, it follows that

$$\omega_\infty = 2 \lim_{B \to \infty} \frac{1}{B \log B} \int_{\mathcal{D}} \frac{\mathrm{d}u\, \mathrm{d}v\, \mathrm{d}t\, \mathrm{d}z}{2\sqrt{t^2 LC(u, v) - z^2}},$$

where we have set $LC(u, v) = L(u, v)C(u, v)$ and $\mathcal{D}$ is the set of $(u, v, t, z) \in \mathbb{R}^4$ such that

$$0 < \max\{u^2, v^2\}t \leqslant B, \quad 0 < z < t\sqrt{LC(u, v)}, \quad 1 \leqslant t \leqslant B, \quad LC(u, v) > 0.$$

In view of the familiar formula

$$\int_0^{\sqrt{S}} \frac{\mathrm{d}s}{\sqrt{S-s^2}} = \frac{\pi}{2},$$

it readily follows that

$$\omega_\infty = \frac{\pi}{2} \sum_{\substack{\varepsilon_1,\varepsilon_2 \in \{\pm 1\} \\ \varepsilon_1\varepsilon_2 = 1}} \mathrm{vol}(R^{\varepsilon_1,\varepsilon_2}(1)).$$

Turning to the $p$-adic densities, we have

$$\omega_p = \lim_{n\to\infty} p^{-4n} \left\{(y,z,t,u,v) \in \mathcal{T}(\mathbb{Z}/p^n\mathbb{Z}) : p \nmid (u,v), p \nmid (y,z,t)\right\}.$$

Recall the definition of $\mathcal{E}$ from §1 and the identities [2, eqs. (2.3) and (2.5)]. To calculate $\omega_2$ we observe that $t$ is odd in any solution to be counted. Since there are $2^{n-1}$ odd integers in the interval $[1, 2^n]$ it follows that

$$\begin{aligned}\omega_2 &= \lim_{n\to\infty} 2^{-3n-1} \# \left\{ (u,v,y,z) \in (\mathbb{Z}/2^n\mathbb{Z})^4 : \begin{array}{l} LC(u,v) \equiv y^2+z^2 \,(\mathrm{mod}\, 2^n), \\ 2 \nmid (u,v) \end{array} \right\} \\ &= \lim_{n\to\infty} 2^{-2n} \# \left\{(u,v) \in (\mathbb{Z}/2^n\mathbb{Z})^2 : LC(u,v) \in \mathcal{E} \,(\mathrm{mod}\, 2^n),\ 2 \nmid (u,v)\right\}.\end{aligned}$$

For any binary form $F \in \mathbb{Z}[u,v]$ and prime power $p^e$, let

$$\widetilde{\varrho}_F(p^e) = p^{-2(e+1)} \# \left\{(u,v) \in (\mathbb{Z}/p^{e+1}\mathbb{Z})^2 :\ p^e \mid F(u,v),\ p \nmid (u,v)\right\}. \tag{7.3}$$

Suppose now that $p \equiv 3 \,(\mathrm{mod}\, 4)$. Then we obtain

$$\begin{aligned}\omega_p &= \lim_{n\to\infty} \frac{1-\frac{1}{p}}{p^{3n}} \# \left\{ (u,v,y,z) \in (\mathbb{Z}/p^n\mathbb{Z})^4 : \begin{array}{l} LC(u,v) \equiv y^2+z^2 \,(\mathrm{mod}\, p^n), \\ p \nmid (u,v) \end{array} \right\} \\ &= \Big(1-\frac{1}{p^2}\Big) \sum_{\nu \geqslant 0} (-1)^\nu \widetilde{\varrho}_{LC}(p^\nu).\end{aligned}$$

Finally, when $p \equiv 1 \,(\mathrm{mod}\, 4)$, we break the cardinality according to the value of $v_p(t)$. It follows that

$$\omega_p = 1 - \frac{1}{p^2} + \Big(1-\frac{1}{p}\Big)^2 \sum_{\nu \geqslant 1} \widetilde{\varrho}_{LC}(p^\nu),$$

in this case.

We now return to our expression (7.2) for $c_0$. Carrying out the summation over $\ell$, we find that

$$\sum_{\ell=1}^{\infty} \frac{\mu(\ell)}{\ell^2} W^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k},\ell) = \widetilde{W}^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k}) = \prod_p \widetilde{W}_p^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k}),$$

for suitable factors $\widetilde{W}_p^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k})$. In view of (6.10), one has

$$\widetilde{W}_2^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k}) = 4\lim_{n\to\infty} 2^{-2n}\#\left\{\mathbf{x}\in(\mathbb{Z}/2^n\mathbb{Z})^2 : \begin{array}{l} L(\mathbf{x})\in\varepsilon_1 d_3\mathcal{E}\,(\mathrm{mod}\,2^n),\\ C(\mathbf{x})\in\varepsilon_1 d_3\mathcal{E}\,(\mathrm{mod}\,2^n),\\ 2\nmid \mathbf{x}\end{array}\right\}.$$

It is clear that for any $\mathbf{x}$ counted here we have both $LC(\mathbf{x})\in\mathcal{E}\,(\mathrm{mod}\,2^n)$ and $LC(-\mathbf{x})\in\mathcal{E}\,(\mathrm{mod}\,2^n)$. Conversely, if $\mathbf{x}\in(\mathbb{Z}/2^n\mathbb{Z})^2$ satisfies $LC(\mathbf{x})\in\mathcal{E}\,(\mathrm{mod}\,2^n)$, then either $L(\mathbf{x})\in\varepsilon_1 d_3\mathcal{E}\,(\mathrm{mod}\,2^n)$ or $L(\mathbf{x})\in-\varepsilon_1 d_3\mathcal{E}\,(\mathrm{mod}\,2^n)$. In this way we conclude that

$$\widetilde{W}_2^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k}) = 2\omega_2,$$

in the above notation. Next, when $p>2$ we deduce from (6.9) that

$$\widetilde{W}_p^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k}) = \Big(1-\frac{\chi(p)}{p}\Big)^2\sum_{\nu_1,\nu_2\geqslant 0}\chi(p^{\nu_1+\nu_2})\widetilde{\varrho}(p^{N_1},p^{N_2}),$$

with

$$\begin{aligned}&\widetilde{\varrho}(p^{N_1},p^{N_2})\\ &\quad = p^{-2(N_1+N_2+1)}\#\left\{\mathbf{x}\in(\mathbb{Z}/p^{N_1+N_2+1}\mathbb{Z})^2 : \begin{array}{l} p^{N_1}\mid L(\mathbf{x}),\ p^{N_2}\mid C(\mathbf{x}),\\ p\nmid\mathbf{x}\end{array}\right\}.\end{aligned}$$

Thus $\widetilde{W}_p^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k})$ is independent of $\varepsilon_1,\varepsilon_2$ and so $\widetilde{W}_p^{\varepsilon_1,\varepsilon_2}(\mathbf{d},\mathbf{k}) = \widetilde{W}_p(\mathbf{d},\mathbf{k})$, say.

An easy calculation reveals that

$$\prod_p \frac{1-\frac{\chi(p)}{p}}{1+\frac{\chi(p)}{p}} = \frac{4}{\pi}\cdot\frac{\pi}{2} = 2.$$

Our work so far has therefore shown that $c_0 = \omega_\infty\omega_2\tau$, with

$$\begin{aligned}\tau = \sum_{d\in\mathfrak{D}}\frac{\mu(d)r_0(d)\varphi^\dagger(d)}{d}&\sum_{\substack{\mathbf{d}\in\mathbb{N}^3\\ d=d_1d_2\\ d_3\mid\Delta}}\chi(d_3)\mu(d_3)\\ &\times\sum_{kk'\mid\gcd(\Delta,d)}\frac{\mu(k')}{2^{\omega(k)}}\prod_{p>2}\Big(\frac{1+\frac{\chi(p)}{p}}{1-\frac{\chi(p)}{p}}\Big)\widetilde{W}_p(\mathbf{d},\mathbf{k}).\end{aligned}$$

We may write $\tau=\prod_{p>2}\tau_p$. Our final task in this paper is to show that $\tau_p=\omega_p$ for each odd prime $p$.

Let $\alpha = v_p(\Delta)$. We will deal here only with the harder case $\alpha \geqslant 1$, the case $\alpha = 0$ being an easy modification. Suppose that $p \equiv 3 \pmod 4$. In this case it is clear that

$$\begin{aligned}\tau_p &= \Big(\frac{1-\frac{1}{p}}{1+\frac{1}{p}}\Big)\Big(1+\frac{1}{p}\Big)^2 \sum_{\nu_1,\nu_2\geqslant 0} \sum_{0\leqslant\delta_3\leqslant 1} (-1)^{\nu_1+\nu_2}\widetilde{\varrho}(p^{\nu_1+\delta_3}, p^{\nu_2+\delta_3}) \\ &= \Big(1-\frac{1}{p^2}\Big) \sum_{\mu_1,\mu_2\geqslant 0} \sum_{0\leqslant\delta_3\leqslant 1} \overline{\varrho}(p^{2\mu_1+\delta_3}, p^{2\mu_2+\delta_3}),\end{aligned}$$

where

$$\overline{\varrho}(p^{n_1}, p^{n_2}) = p^{-2(n_1+n_2+1)} \# \left\{ (u,v) \in (\mathbb{Z}/p^{n_1+n_2+1}\mathbb{Z})^2 : \begin{array}{l} p^{n_1} \| L(u,v), \\ p^{n_2} \| C(u,v), \\ p \nmid (u,v) \end{array} \right\}.$$

Setting $\overline{\varrho}(p^n)$ for the analogous density in which one has $p^n \| LC(u,v)$ instead of the pair of conditions present in $\overline{\varrho}(p^{n_1}, p^{n_2})$, one finds that

$$\tau_p = \Big(1-\frac{1}{p^2}\Big) \sum_{\mu\geqslant 0} \overline{\varrho}(p^{2\mu}) = \omega_p,$$

as required.

Suppose now that $p \equiv 1 \pmod 4$. Then we have

$$\tau_p = \Big(1-\frac{1}{p^2}\Big) \sum_{0\leqslant\delta\leqslant 1} \frac{(-1)^\delta r_0(p^\delta)\varphi^\dagger(p^\delta)}{p^\delta} \sum_{\substack{\delta_1,\delta_2,\delta_3\in\{0,1\} \\ \delta_1+\delta_2=\delta}} (-1)^{\delta_3} f_p(\delta_1,\delta_2,\delta_3)$$

with

$$f_p(\delta_1,\delta_2,\delta_3) = \sum_{\nu_1,\nu_2\geqslant 0} \sum_{\substack{\kappa,\kappa'\geqslant 0 \\ \kappa+\kappa'\leqslant\delta}} \frac{(-1)^{\kappa'}}{2^\kappa} \widetilde{\varrho}(p^{N_1}, p^{N_2})$$

and $N_i = \max\{\kappa+\kappa', \nu_i+\delta_i+\delta_3\}$ for $i = 1, 2$. We claim that

$$f_p(\delta_1,\delta_2,\delta_3) = \sum_{\nu_1,\nu_2\geqslant 0} \frac{(\nu_1+1)(\nu_2+1)}{2^{\min\{\delta, N_1', N_2'\}}} \overline{\varrho}(p^{N_1'}, p^{N_2'}), \tag{7.4}$$

with $N_i' = \nu_i + \delta_i + \delta_3$ for $i = 1, 2$. We begin by noting that

$$f_p(\delta_1,\delta_2,\delta_3) = \sum_{\nu_1,\nu_2\geqslant 0} \sum_{\substack{\kappa,\kappa'\geqslant 0 \\ \kappa+\kappa'\leqslant\min\{\delta, N_1', N_2'\}}} \frac{(-1)^{\kappa'}}{2^\kappa} (\nu_1+1)(\nu_2+1) \overline{\varrho}(p^{N_1'}, p^{N_2'}).$$

But it is clear that

$$\sum_{\substack{\kappa,\kappa'\geqslant 0\\ \kappa+\kappa'\leqslant \min\{\delta,N_1',N_2'\}}} \frac{(-1)^{\kappa'}}{2^{\kappa}} = \frac{1}{2^{\min\{\delta,N_1',N_2'\}}},$$

from which the claim follows.

Given (7.4) we are now led to consider the quantity

$$f_p(\delta) = \sum_{\substack{\delta_1,\delta_2,\delta_3\in\{0,1\}\\ \delta_1+\delta_2=\delta}} (-1)^{\delta_3} \sum_{\nu_1,\nu_2\geqslant 0} \frac{(\nu_1+1)(\nu_2+1)}{2^{\min\{\delta,N_1',N_2'\}}} \overline{\varrho}(p^{N_1'},p^{N_2'}),$$

for each $\delta\in\{0,1\}$. Let $N_i''=\nu_i+\delta_i$ for $i=1,2$. We may write

$$\begin{aligned} f_p(\delta) &= \sum_{\substack{\delta_1,\delta_2\geqslant 0\\ \delta_1+\delta_2=\delta}} \sum_{\nu_1,\nu_2\geqslant 0} \big((\nu_1+1)(\nu_2+1)-\nu_1\nu_2\big) \frac{\overline{\varrho}(p^{N_1''},p^{N_2''})}{2^{\min\{\delta,N_1'',N_2''\}}}\\ &= \sum_{\substack{\delta_1,\delta_2\geqslant 0\\ \delta_1+\delta_2=\delta}} \sum_{\nu_1,\nu_2\geqslant 0} (\nu_1+\nu_2+1) \frac{\overline{\varrho}(p^{N_1''},p^{N_2''})}{2^{\min\{\delta,N_1'',N_2''\}}}. \end{aligned}$$

When $\delta=1$ and

$$\min\{1,v_p(L(\mathbf{x})),v_p(C(\mathbf{x}))\} = \min\{1,N_1'',N_2''\}\geqslant 1,$$

with $p^{\nu+\delta}\|LC(\mathbf{x})$, there are two choices of $(\delta_1,\delta_2)$ such that $\delta_1+\delta_2=\delta$, $p^{\nu_1+\delta_1}\|L(\mathbf{x})$, $p^{\nu_2+\delta_2}\|C(\mathbf{x})$ and $\nu=\nu_1+\nu_2$. Thus

$$f_p(\delta) = \sum_{\nu\geqslant 0}(\nu+1)\overline{\varrho}(\nu+\delta) = \sum_{\nu\geqslant 0}\widetilde{\varrho}_{LC}(p^{\nu+\delta}),$$

in this case. The same is true when $\delta=0$. Recalling that $p^{-1}\varphi^{\dagger}(p)=(p+1)^{-1}$, we deduce that

$$\begin{aligned} \tau_p &= \Big(1-\frac{1}{p^2}\Big) \sum_{0\leqslant\delta\leqslant 1} \frac{(-1)^{\delta} r_0(p^{\delta})\varphi^{\dagger}(p^{\delta})}{p^{\delta}} f_p(\delta)\\ &= \Big(1-\frac{1}{p^2}\Big)\Big\{1+\sum_{\nu\geqslant 1}\widetilde{\varrho}_{LC}(p^{\nu})\Big(1-\frac{2}{p+1}\Big)\Big\}\\ &= \omega_p. \end{aligned}$$

This completes the proof that the value of the leading constant in Theorem 1 agrees with the prediction of Peyre.

## References


[1] R. de la Bretèche and T. D. Browning, *Sums of arithmetic functions over values of binary forms,* Acta Arithmetica **125** (2007), 291–304.

[2] R. de la Bretèche and T. D. Browning, *Binary linear forms as sums of two squares,* Compositio Mathematica **144** (2008), 1375–1402.

[3] R. de la Bretèche and T. D. Browning, *Le problème des diviseurs pour des formes binaires de degré 4,* Journal für die Reine und Angewandte Mathematik **646** (2010), 1–44.

[4] R. de la Bretèche, T. D. Browning and E. Peyre, *On Manin's conjecture for a family of Châtelet surfaces,* Annals of Mathematics, to appear.

[5] T. D. Browning, *Linear growth for Châtelet surfaces,* Mathematische Annalen **346** (2010), 41–50.

[6] J.-L. Colliot-Thélène, J.-J. Sansuc and P. Swinnerton-Dyer, *Intersections of two quadrics and Châtelet surfaces. I,* Journal für die Reine und Angewandte Mathematik **373** (1987), 37–107.

[7] J.-L. Colliot-Thélène, J.-J. Sansuc and P. Swinnerton-Dyer, *Intersections of two quadrics and Châtelet surfaces. II,* Journal für die Reine und Angewandte Mathematik **374** (1987), 72–168.

[8] S. Daniel, *On the divisor-sum problem for binary forms,* Journal für die Reine und Angewandte Mathematik **507** (1999), 107–129.

[9] H. Davenport, *Cubic forms in 16 variables,* Proceedings of the Royal Society of Edinburgh, Section A **272** (1963), 285–303.

[10] R. Dedekind, *Gesammelte mathematische Werke*, Band **1**, Vieweg & Sohn, Braunschweig, 1930.

[11] J. Franke, Y. I. Manin and Y. Tschinkel, *Rational points of bounded height on Fano varieties,* Inventiones Mathematicae **95** (1989), 421–435.

[12] D. R. Heath-Brown, *Linear relations amongst sums of two squares,* in *Number Theory and Algebraic Geometry*, London Mathematical Society Lecture Note Series **303**, Cambridge University Press, 2003, pp. 133–176.

[13] H. Heilbronn, *Zeta-functions and L-functions,* in *Algebraic Number Theory (Proc. Instructional Conf., Brighton, 1965)*, Thompson, Washington, DC, 1967, pp. 204–230.

[14] M. N. Huxley, *A note on polynomial congruences, in Recent Progress in Analytic Number Theory* **1***, (Durham, 1979)*, Academic Press, New York, 1981, pp. 193–196.

[15] H. Iwaniec and R. Munshi, *Cubic polynomials and quadratic forms,* Journal of the London Mathematical Society **81** (2010), 45–64.

[16] G. Marasingha, *Almost primes represented by binary forms,* Journal of the London Mathematical Society **82** (2010), 295–316.

[17] E. Peyre, *Hauteurs et nombres de Tamagawa sur les variétés de Fano,* Duke Mathematical Journal **79** (1995), 101–218.

[18] C. L. Stewart, *On the number of solutions of polynomial congruences and Thue equations,* Journal of the American Mathematical Society **4** (1991), 793–835.