BINARY QUARTIC FORMS WITH VANISHING J-INVARIANT

STANLEY YAO XIAO

Abstract. We obtain an asymptotic formula for the number of $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of irreducible, totally real binary quartic forms with integer coefficients with vanishing J-invariant. These results give a case where one is able to count integral orbits inside a relatively open real orbit of a variety of degree at least three which is closed under a group action. As a consequence, we give an asymptotic formula for the number of $\mathrm{GL}_2(\mathbb{Z})$-classes of irreducible binary quartic forms with vanishing J-invariant and Galois group C_4, ordered by discriminant. Our method of proof introduces a new observation regarding taking powers in the class group of quadratic forms of a given discriminant and lattices associated to representable primes (see appendix by Erick Knight).

1. Introduction

Let

$$F(x,y) = a_4x^4 + a_3x^3y + a_2x^2y^2 + a_1xy^3 + a_0y^4 \in \mathbb{R}[x,y] \tag{1.1}$$

be a binary quartic form, and put $V_4(\mathbb{R})$ for the 5-dimensional vector space of real binary quartic forms. The group $\mathrm{GL}_2(\mathbb{R})$ acts on $V_4(\mathbb{R})$ via the *substitution action*, defined for $T = \left(\begin{smallmatrix} t_1 & t_2 \\ t_3 & t_4 \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbb{R})$ and $F \in V_4(\mathbb{R})$ by

$$F_T(x,y) = F(t_1x + t_2y, t_3x + t_4y). \tag{1.2}$$

It is well-known that the ring of relative polynomial invariants of the substitution action of $\mathrm{GL}_2(\mathbb{R})$ on $V_4(\mathbb{R})$ is a polynomial ring generated by two elements, commonly denoted as I and J. They are given by

$$I(F) = 12a_4a_0 - 3a_3a_1 + a_2^2 \tag{1.3}$$

and

$$J(F) = 72a_4a_2a_0 + 9a_3a_2a_1 - 27a_4a_1^2 - 27a_0a_3^2 - 2a_2^3. \tag{1.4}$$

Both the quadric defined by $I(F) = 0$ and the cubic defined by $J(F) = 0$ are invariant under $\mathrm{GL}_2(\mathbb{R})$; that is, for all $F \in V_{\mathbb{R}}$ and $T \in \mathrm{GL}_2(\mathbb{R})$, we have $J(F) = 0$ if and only if $J(F_T) = 0$.

Put

$$\mathcal{V}_4(\mathbb{R}) = \{F \in V_4(\mathbb{R}) : J(F) = 0\} \tag{1.5}$$

for the real cubic threefold defined by the vanishing of J in $V_4(\mathbb{R})$. $\mathcal{V}_4(\mathbb{R}) \setminus \{F : \Delta(F) = 0\}$ consists of three relatively open orbits under the substitution action of $\mathrm{GL}_2(\mathbb{R})$, consisting of non-singular forms with 0, 2, or 4 real linear factors respectively. We shall denote by $\mathcal{V}_4^{(i)}(\mathbb{R})$ the orbit of $\mathcal{V}_4(\mathbb{R})$ consisting of forms with $4 - 2i$ real linear factors. In particular,

$$\mathcal{V}_4^{(0)}(\mathbb{R}) = \{F \in \mathcal{V}_4(\mathbb{R}) : F \text{ has 4 real linear factors}\}.$$

The *discriminant* $\Delta(F)$ of a binary quartic form F is expressible in terms of I and J as

$$\Delta(F) = \frac{4I(F)^3 - J(F)^2}{27}. \tag{1.6}$$

Put

$$W_n(\mathbb{Z}) = \{\mathrm{GL}_2(\mathbb{Z})\text{-orbits of integral binary } n\text{-ic forms}\}.$$

Since $I(F), J(F)$ are $\mathrm{GL}_2(\mathbb{Z})$-invariants, for any class $w \in W_4(\mathbb{Z})$ and any $F, G \in w$ we have $I(F) = I(G)$ and $J(F) = J(G)$. Thus, the values of I, J are well-defined on the class w. Now put

$$\mathcal{W}_4(\mathbb{Z}) = \{w \in W_4(\mathbb{Z}) : J(w) = 0\}. \tag{1.7}$$

The main goal of this paper is to establish asymptotic formulae for two subclasses of $\mathcal{W}_4(\mathbb{Z})$ with non-zero discriminant, where we count the orbits by discriminant. This count is a priori finite by a well-known result

of Borel and Harish-Chandra [7]. Since the number of real linear factors are preserved under $\mathrm{GL}_2(\mathbb{R})$ action, one can define the number of real linear factors for orbits in $\mathcal{W}_4(\mathbb{R})$. Indeed, we shall put

$$\mathcal{W}_4^{(i)}(\mathbb{Z}) = \{w \in \mathcal{W}_4(\mathbb{Z}) : F \in \mathcal{V}_4^{(i)}(\mathbb{R}) \text{ for all } F \in w\}.$$

The first family we shall consider is $\mathcal{W}_4^{(0)}(\mathbb{Z})$. For a positive number X, put

$$N(X) = \#\{w \in \mathcal{W}_4^{(0)}(\mathbb{Z}) : \Delta(w) \neq 0, \Delta(w) \leq X\}. \tag{1.8}$$

We denote the set on the right hand side above by $\mathcal{W}_4^{(0,\dagger)}(X)$. Further note that for all binary quartic forms F with real coefficients and 4 real linear factors, we have $\Delta(F) > 0$.

The forms with vanishing J-invariant can be characterized by the fact that their *Hessian* covariants are perfect squares in $\mathbb{C}[x, y]$. The Hessian covariant of F, denoted as H_F, is given by

$$\begin{aligned} H_F(x,y) = (3a_3^2 - 8a_4a_2)x^4 + 4(a_3a_2 - 6a_4a_1)x^3y + 2(2a_2^2 - 24a_4a_0 - 3a_3a_1)x^2y^2 \\ +4(a_2a_1 - 6a_3a_0)xy^3 + (3a_1^2 - 8a_2a_0)y^4. \end{aligned} \tag{1.9}$$

Then $F \in \mathcal{V}_4(\mathbb{R})$ if and only if there exists a quadratic form f with complex coefficients such that $f^2 | H_F$ as elements in $\mathbb{C}[x, y]$. Further, one can take f to be a form with co-prime integer coefficients and non-zero discriminant when $F \in \mathcal{V}_4(\mathbb{Z})$ and $I(F) \neq 0$. Moreover when $F \in \mathcal{V}_4^{(0)}(\mathbb{R})$ we may take f to have real coefficients and $\Delta(f) < 0$ (see Lemma 5.1). Observe that if H_F is divisible by the square of a reducible quadratic form, then so will H_{F_T} for any $T \in \mathrm{GL}_2(\mathbb{Z})$. The next family we shall consider will be:

$$\begin{aligned} \mathcal{W}_4^{\star}(\mathbb{Z}) = \{w \in \mathcal{W}_4(\mathbb{Z}) : \text{for all } F \in w, H_F \text{ is divisible by the square of a reducible} \\ \text{quadratic form } f\}. \end{aligned}$$

We now put

$$M(X) = \#\{w \in \mathcal{W}_4^{\star}(\mathbb{Z}) : |\Delta(w)| \leq X, w \text{ is irreducible.}\}$$

The main theorems of our paper are the following counting results for $N(X)$ and $M(X)$:

Theorem 1.1. *The asymptotic formula*

$$N(X) = \frac{6\sqrt[3]{2}\zeta(2)}{7\zeta(3)} X^{1/3} \log X + O\left(X^{1/3}\right).$$

holds.

Theorem 1.2. *The asymptotic formula*

$$M(X) = \frac{\zeta(2)}{6\sqrt[3]{4}\zeta(3)} X^{1/3} \log X + O\left(X^{1/3}\right)$$

holds.

We note that the methods introduced in this paper just barely fall short of being able to give the analogous result in Theorem 1.1 for $\mathcal{W}_4^{(2)}(\mathbb{Z})$ and $\mathcal{W}_4^{(4)}(\mathbb{Z})$. This lacuna will be filled in in a future paper.

Since by definition elements in $\mathcal{W}_4^{(0)}$ we have $J(w) = 0$, and thus for any such orbit and any irreducible $F \in w$, we have that the Galois group of the splitting field of F is a subgroup of D_4 (see [20] for a full treatment). We have the following:

Theorem 1.3. *Let f be a positive definite, primitive integral binary quadratic form. Let $F \in \mathcal{V}_f(\mathbb{Z})$ be an irreducible binary quartic form. Then* $\mathrm{Gal}(F) \cong C_4$ *if and only if $-\Delta(f)$ is a square.*

The proof of Theorem 1.3 relies on the fact that whenever $-\Delta(f) = \square$, any form $F \in \mathcal{V}_f(\mathbb{Z})$ with square discriminant is *necessarily reducible*. Thus, following the criteria determining all possible Galois groups of quartic forms in [9], all irreducible elements must necessarily have Galois group C_4. This fact is based on the existence of a natural involution on $\mathcal{V}_4^{(0)}(\mathbb{R})$.

Let $F \in V_{\mathbb{R}}$ be as given in (1.1). It has a natural *sextic covariant* given by the Jacobian determinant of F and the Hessian covariant H_F of F. It has the following explicit formula:

$$\begin{aligned} F_6(x,y) = &(a_3^3 + 8a_4^2a_1 - 4a_4a_3a_2)x^6 + 2(16a_4^2a_0 + 2a_4a_3a_1 - 4a_4a_2^2 + a_3^2a_2)x^5y \\ &+ 5(8a_4a_3a_0 + a_3^2a_1 - 4a_4a_2a_1)x^4y^2 + 20(a_3^2a_0 - a_4a_1^2)x^3y^3 \\ &- 5(8a_4a_1a_0 + a_3a_1^2 - 4a_3a_2a_0)x^2y^4 - 2(16a_4a_0^2 + 2a_3a_1a_0 - 4a_2^2a_0 + a_2a_1^2)xy^5 \\ &- (a_1^3 + 8a_3a_0^2 - 4a_2a_1a_0)y^6. \end{aligned} \tag{1.10}$$

In [24] we proved that F_6 is always a *Klein form* (see [1]). Moreover, when $F \in \mathbb{R}[x,y]$ and $J(F) = 0$ it admits a factorization of the shape

$$F_6(x,y) = f(x,y)G_F(x,y),$$

where $f(x,y)$ is a binary quadratic form with real coefficients such that $f^2|H_F$ over $\mathbb{C}$ and $J(G_F) = 0$. Recall that when F is totally real we will see that f has negative discriminant. We now choose $f = f_F$ such that $\Delta(f) = -4$, and define G_F as the quotient $F_6/f \in \mathbb{R}[x,y]$. We then have the following:

Theorem 1.4. *The map $\Xi : \mathcal{V}_4^{(0)}(\mathbb{R}) \to V_{\mathbb{R}}$ defined by*

$$\Xi(F) = G_F$$

satisfies $\Xi^2(F) = \alpha_F F$ for some $\alpha_F \in \mathbb{R}$.

Of course the involution Ξ need not restrict to an involution from $\mathcal{V}_4^{(0)}(\mathbb{Z})$ to $\mathcal{V}_4^{(0)}(\mathbb{Z})$. However we will see that there is a rational version $\Xi_{\mathbb{Q}}$ of Ξ such that $\Xi_{\mathbb{Q}}^2(F) = cF$, for some $c \in \mathbb{Z}$. This is enough to control the reducible forms $F \in \mathcal{V}_f(\mathbb{Z})$ with $\Delta(F) = \square$ or $-\Delta(F) = \square$.

Theorem 1.3 and an easier case of Theorem 1.1 have the following attractive consequence. For $w \in W_4(\mathbb{Z})$ define the *Bhargava-Shankar height* to be

$$H_{\mathrm{BS}}(w) = \max\{|I(w)|^3, J(w)^2/4\}. \tag{1.11}$$

For a transitive subgroup G of the symmetric group S_4, put

$$\mathcal{N}_G(X) = \#\{w \in W_4(\mathbb{Z}) : H_{\mathrm{BS}}(w) \le X, \mathrm{Gal}(w) \cong G\}$$

and

$$\mathcal{M}_G(X) = \#\{w \in W_4(\mathbb{Z}) : |\Delta(w)| \le X, \mathrm{Gal}(w) \cong G\}.$$

Let $N_G(X), M_G(X)$ denote respectively the sub-count of $\mathcal{N}_G(X), \mathcal{M}_G(X)$ restricted to orbits with $J = 0$. We obtain the theorem:

Theorem 1.5. *Let $\varepsilon > 0$. We have the asymptotic formulae*

$$N_{C_4}(X) = \frac{7}{9}X^{1/3} + O_\varepsilon\left(X^{1/3-\varepsilon}\right)$$

and

$$M_{C_4}(X) = \frac{7}{6\sqrt[3]{2}}X^{1/3} + O_\varepsilon\left(X^{1/3-\varepsilon}\right).$$

Theorem 1.5 then implies:

Corollary 1.6. *There exist positive numbers c_0, c_1, c_2 such that for any $X > c_0$ we have*

$$\mathcal{N}_{C_4}(X) > c_1X^{1/3}$$

and

$$\mathcal{M}_{C_4}(X) > c_2X^{1/3}.$$

In [20] we proved, along with Tsang, that the number of C_4-forms with a fixed Cremona covariant and bounded Bhargava-Shankar height X is $O_{f,\varepsilon}\left(X^{1/6+\varepsilon}\right)$. In fact the forms corresponding to C_4-forms in a fixed family parametrized by the quadratic form f are the rational points on a certain toric, singular del Pezzo surface of degree 4. This leads us to conjecture the following:

Conjecture 1.7. *Let $\varepsilon > 0$. Then*

$$\mathcal{M}_{C_4}(X) = M_{C_4}(X) + O_\varepsilon\left(X^{1/6+\varepsilon}\right) \text{ and } \mathcal{N}_{C_4}(X) = N_{C_4}(X) + O_\varepsilon\left(X^{1/6+\varepsilon}\right).$$

In [3], Bhargava and Shankar proved that

$$\mathcal{N}_{S_4}(X) = \frac{44\zeta(2)}{135}X^{5/6} + O_\varepsilon\left(X^{3/4+\varepsilon}\right),$$

which they used to obtain their magnificent theorem on the boundedness of average Mordell-Weil rank of elliptic curves over $\mathbb{Q}$. It remains a significant challenge to estimate $\mathcal{M}_{S_4}(X)$ from above.

In [20] and [21] Tsang and I proved that

$$\mathcal{N}_{D_4}(X) \gg X^{1/2}\log X \tag{1.12}$$

and

$$\mathcal{M}_{D_4}(X) \gg X^{1/2}(\log X)^2. \tag{1.13}$$

We also showed that

$$\mathcal{N}_{V_4}(X), \mathcal{M}_{V_4}(X) \gg X^{1/3}. \tag{1.14}$$

We expect that both (1.12) and (1.13) represent the true orders of magnitude. Comparing (1.14) and Corollary 1.6, one has to wonder whether V_4-forms or C_4-forms are more numerous.

It remains a difficult challenge to estimate $\mathcal{N}_{A_4}(X), \mathcal{M}_{A_4}(X)$.

Theorem 1.1 represents the first case where one can count integral $G(\mathbb{Z})$-orbits inside a relatively open orbit $G(\mathbb{R})\cdot v$, where $v \in V(\mathbb{R})$ sits inside a *proper subvariety* of degree at least three which is closed under the action of $G(\mathbb{R})$. In our case, the group G is $\mathrm{GL}_2(\mathbb{R})$ and the variety is the cubic threefold in $V_4(\mathbb{R})$ given by $J(F) = 0$. The methods we employ in this paper, while heavily inspired by the work of Bhargava, do not directly involve his geometry of numbers method and the action of $\mathrm{GL}_2(\mathbb{R})$ on $\mathcal{V}_4(\mathbb{R})$ is not directly exploited. Instead, we partition $\mathcal{V}_4(\mathbb{R})$ into families indexed by $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of integral binary quadratic forms, as we did in [20]. This reduces the problem of counting integral orbits in $\mathcal{V}_4(\mathbb{R})$ to counting integer points, sorted by discriminant, inside a countable collection of 2-dimensional vector spaces inside $\mathcal{V}_4(\mathbb{R})$. We then use a wide assortment of results regarding binary quadratic forms to help establish Theorem 1.1. Of particular note is Proposition 5.9, which is a novel observation regarding the change in the $\mathrm{SL}_2(\mathbb{Z})$-class of quadratic forms as one performs 'Hensel lifting' of lattices containing primitive solutions to the congruence $f(x, y) \equiv 0 \pmod{p^k}$.

We remark that S. Ruth, in his thesis, gave an argument that essentially counts the number of $\mathrm{GL}_2(\mathbb{Z})$-orbits of quartic Klein forms (see [1] for a modern treatment), that is, those quartic forms with $I(F) = 0$. However there is some doubt that his application of Heath-Brown's circle method [14] is acceptable. In January 2019 the author heard a lecture given by A. Alpoge on this matter, which resolved the issue by giving an argument which circumvents the problematic application of Heath-Brown's method. Therefore, the asymptotic formula given for the number of quartic Klein forms of bounded discriminant given by Ruth is correct.

Theorem 1.2, in comparison, is relatively straightforward. This is mostly because the class number of reducible quadratic forms is very easy to understand, and that the set of discriminants of reducible quadratic forms is equal to the set of square integers, which is a very thin set. We give the proof of Theorem 1.2 in Section 9.

In view of Theorems 1.1 and 1.2, all that is needed to prove the full asymptotic formula for the number of $\mathrm{GL}_2(\mathbb{Z})$-orbits of binary quartic forms F with vanishing J-invariant is to count the number of integral orbits whose Hessians are divisible by the square of an *irreducible, indefinite* binary quadratic form. There are significant barriers to carrying out the arguments in this paper to handle this case, but there is another method to count such orbits. We wish to expand on this in future work.

Finally, Proposition 5.9 appears to be a new observation regarding taking powers in the class group of quadratic forms of a given negative discriminant and may be of separate interest. The author thanks Erick Knight for providing the proof in the appendix which is much more elegant than his own.

Acknowledgements. We thank an anonymous referee who identified a key error in a previous version of this paper. We thank J. Friedlander for suggesting the paper [17], which contains the key ideas necessary to carry out the proof of Theorem 1.1. We thank A. Alpoge for his lecture given at the Joint Mathematics Meetings in Baltimore, which settled the question of whether the correct asymptotic formula for the number of $\mathrm{GL}_2(\mathbb{Z})$-orbits of quartic Klein forms has been obtained.

2. Parametrizing quartic forms with $J = 0$ by the Hessian

In this section, we will refine our parametrization theorem in our work with Tsang in [20] to provide a parametrization theorem for binary quartic forms with vanishing J-invariant. For a binary quadratic form f with integer coefficients, put $\mathcal{C}(f)$ for its $\mathrm{GL}_2(\mathbb{Z})$-equivalence class, and put

$$\mathcal{V}_f(\mathbb{R}) = \{F \in \mathcal{V}_4(\mathbb{R}) : f^2 | H_F\},$$

and $\mathcal{V}_f(\mathbb{Z})$ for the subset of $\mathcal{V}_f(\mathbb{R})$ consisting of those forms with integer coefficients. We then put

$$\mathcal{W}_f(\mathbb{Z}) = \{w \in \mathcal{W}_4(\mathbb{Z}) : \exists F \in w \text{ s.t. } f^2 | H_F\}$$

and $\mathcal{W}_f^\dagger(\mathbb{Z})$ to be the subset consisting of irreducible elements. Notice that if f and g are $\mathrm{GL}_2(\mathbb{Z})$-equivalent, then $\mathcal{W}_f(\mathbb{Z}) = \mathcal{W}_g(\mathbb{Z})$; hence $\mathcal{W}_f(\mathbb{Z})$ only depends on $\mathcal{C}(f)$, so we write $\mathcal{W}_{\mathcal{C}(f)}(\mathbb{Z})$ instead. We then have the following result:

Proposition 2.1. *We have that $\mathcal{W}_4(\mathbb{Z})$ is given by the disjoint union*

$$\mathcal{W}_4(\mathbb{Z}) = \bigcup_{\mathcal{C}(f)} \mathcal{W}_{\mathcal{C}(f)}(\mathbb{Z}),$$

where $\mathcal{C}(f)$ varies over all $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of primitive, integral binary quadratic forms with non-zero discriminant.

Our goal is to obtain, for each $\mathcal{C}(f)$, a set of representatives in $\mathcal{V}_f(\mathbb{Z})$ for $\mathcal{W}_{\mathcal{C}(f)}(\mathbb{Z})$. We begin with the following lemma:

Lemma 2.2. *Let f, g be two binary quadratic forms with co-prime integer coefficients and let $F, G \in \mathcal{V}_{\mathbb{Z}}$ be such that $F \in \mathcal{V}_{f,\mathbb{Z}}, G \in \mathcal{V}_{g,\mathbb{Z}}$. If F and G are $\mathrm{GL}_2(\mathbb{Z})$-equivalent, then f is $\mathrm{GL}_2(\mathbb{Z})$-equivalent to either g or $-g$.*

Proof. This follows from the fact that if F, G are $\mathrm{GL}_2(\mathbb{Z})$-equivalent then their Hessians H_F, H_G are $\mathrm{GL}_2(\mathbb{Z})$-equivalent, since H_F is a covariant of F. This implies that f^2 is $\mathrm{GL}_2(\mathbb{Z})$-equivalent to g^2. Taking square roots shows that f is $\mathrm{GL}_2(\mathbb{Z})$-equivalent to $\pm g$. □

We show that $\mathcal{V}_f(\mathbb{Z})$ has a natural structure as a 2-dimensional lattice, and we give an explicit embedding of $\mathcal{V}_f(\mathbb{Z})$ into $\mathbb{Z}^2$ below. First note that from (1.4) we see that $a_2 \equiv 0 \pmod 3$. Put

$$f(x, y) = \alpha x^2 + \beta xy + \gamma y^2, \alpha, \beta, \gamma \in \mathbb{Z}, \tag{2.1}$$

with

$$\mathcal{A}_1 = 4\gamma A - \beta B, \tag{2.2}$$

$$\mathcal{A}_2 = 4\beta\gamma A - (\beta^2 - \alpha\gamma)B, \tag{2.3}$$

and

$$\mathcal{A}_3 = 4\gamma(\beta^2 - \alpha\gamma)A - \beta(\beta^2 - 2\alpha\gamma)B. \tag{2.4}$$

Define the lattice $\mathcal{L}_{f,\alpha}$ by

$$\mathcal{L}_{f,\alpha} = \{(A, B) \in \mathbb{Z}^2 : \mathcal{A}_1 \equiv 0 \pmod{2\alpha}, \mathcal{A}_2 \equiv 0 \pmod{\alpha^2}, \mathcal{A}_3 \equiv 0 \pmod{4\alpha^3}\}. \tag{2.5}$$

We have the following result regarding the family $\mathcal{V}_{f,\mathbb{Z}}$:

Proposition 2.3. *Let $f(x,y) = \alpha x^2 + \beta xy + \gamma y^2, \alpha \neq 0$ be a binary quadratic form with co-prime integer coefficients and non-zero discriminant. The forms in $\mathcal{V}_f(\mathbb{Z})$ are of the form*

$$\left\{ Ax^4 + Bx^3y - \frac{12\gamma A - 3\beta B}{2\alpha}x^2y^2 + \left(\frac{-4\beta\gamma A + (\beta^2 - \alpha\gamma)B}{\alpha^2}\right)xy^3 \right. \\ \left. + \left(\frac{-4\gamma(\beta^2 - \alpha\gamma)A + \beta(\beta^2 - 2\alpha\gamma)B}{4\alpha^3}\right)y^4 : (A,B) \in \mathcal{L}_{f,\alpha} \right\}. \tag{2.6}$$

In particular, the map $\nu : \mathcal{V}_f(\mathbb{Z}) \to \mathcal{L}_{f,\alpha}$ given by $\nu(Ax^4 + Bx^3y + \cdots) = (A,B)$ is a bijection between $\mathcal{V}_f(\mathbb{Z})$ and $\mathcal{L}_{f,\alpha}$.

To prove Proposition 2.3, we shall require some results from [24] and [20] and recall some relevant notions. For a given binary quartic form F with real coefficients, define the *automorphism group* of F (over $\mathbb{R}$) as

$$\mathrm{Aut}_{\mathbb{R}} F = \{T \in \mathrm{GL}_2(\mathbb{R}) : F_T(x,y) = F(x,y)\}. \tag{2.7}$$

For a subring $\mathbb{F}$ of $\mathbb{R}$, define

$$\mathrm{Aut}_{\mathbb{F}} F = \{T \in \mathrm{Aut}_{\mathbb{R}} F : \exists \lambda \in \mathbb{R} \text{ s.t. } \lambda T \in \mathrm{GL}_2(\mathbb{F})\}.$$

For a given binary quadratic form $f(x,y) = \alpha x^2 + \beta xy + \gamma y^2$ with real coefficients, define

$$M_f = \frac{1}{\sqrt{|\Delta(f)|}}\begin{pmatrix} \beta & 2\gamma \\ -2\alpha & -\beta \end{pmatrix}. \tag{2.8}$$

Put

$$V_f(\mathbb{R}) = \{F \in V_4(\mathbb{R}) : M_f \in \mathrm{Aut}_{\mathbb{R}} F\}.$$

We can now prove the following, which identifies $\mathcal{V}_f(\mathbb{R})$ as a plane inside $V_f(\mathbb{R})$:

Lemma 2.4. *Let f be a binary quadratic form with real coefficients and non-zero discriminant. Then $\mathcal{V}_f(\mathbb{R})$ is the plane inside $V_f(\mathbb{R})$ defined by*

$$12\gamma A - 3\beta B + 2\alpha C = 0.$$

Proof. By the results in [20], it follows that for any $F \in V_f(\mathbb{R})$ we have f^2 divides

$$\mathfrak{F}_1(x,y) = \frac{1}{3}\left(H_F(x,y) + 4L_f(F)F(x,y)\right),$$

where

$$L_f(F) = -\frac{12\gamma A - 3\beta B + 2\alpha C}{2\alpha}.$$

It therefore follows that $F \in \mathcal{V}_f(\mathbb{R})$ if and only if $L_f(F) = 0$ or $F(x,y)$ is proportional to H_F. The latter implies that $\Delta(F) = 0$, so the former must hold. □

Proof of Proposition 2.3. By (3.1) of [20], it follows that whenever $F \in V_f(\mathbb{R})$, the xy^3, y^4 coefficients of F are given by linear forms in the coefficients of x^4, x^3y, x^2y^2. In particular, we have

$$F(x,y) = Ax^4 + Bx^3y + Cx^2y^2 + + \left(\frac{4\beta\gamma A - (\beta^2 + 2\alpha\gamma)B + 2\alpha\beta C}{2\alpha^2}\right)xy^3 \\ + \left(\frac{4\gamma(\beta^2 + 2\alpha\gamma)A - \beta(\beta^2 + 4\alpha\gamma)B + 2\alpha\beta^2 C}{8\alpha^3}\right)y^4$$

for $A, B, C \in \mathbb{R}$. Moreover we see that f^2 is proportional to the quartic form

$$\frac{1}{3}\left(H_F(x,y) + 4L_f(F)F(x,y)\right).$$

The condition that $f^2 | H_F$ implies that $L_f(F) = 0$, or equivalently,

$$C = \frac{-12\gamma A + 3\beta B}{2\alpha}.$$

We then see that the condition $F \in V_{\mathbb{Z}}$ is then equivalent to $(A,B) \in \mathcal{L}_{f,\alpha}$, as desired. □

Our aim now is to show that $\mathcal{V}_f(\mathbb{Z})$ is an n-cover for $\mathcal{W}_{\mathcal{C}(f)}(\mathbb{Z})$, where n is a positive integer which is absolutely bounded. Typically we will have $n = 1$. We shall precisely describe when $\mathcal{V}_f(\mathbb{Z})$ fails to be in one-to-one correspondence with $\mathcal{W}_{\mathcal{C}(f)}(\mathbb{Z})$. We shall need the following definition:

Definition 2.5. Let f be a binary quadratic form with integer coefficients and non-zero discriminant. Then f is said to be *ambiguous* if there exists a $\mathrm{GL}_2(\mathbb{Z})$-translate $g = g_2x^2+g_1xy+g_0y^2$ of f such that $g_2|g_1$. We say that f is *opaque* if there exists a $\mathrm{GL}_2(\mathbb{Z})$-translate g of f which takes the shape $g(x,y) = g_2x^2+g_1xy-g_2y^2$.

We summarize some results of [20] as follows:

Proposition 2.6. *Let $f = \alpha x^2 + \beta xy + \gamma y^2$ be a primitive integral binary quadratic form with non-zero discriminant. Then there exists a positive integer n_f such that $\mathcal{V}_f(\mathbb{Z})$ is a n_f-fold cover of $\mathcal{W}_{\mathcal{C}(f)}(\mathbb{Z})$, where*

$$n_f = \begin{cases} 1 & \text{if } f \text{ is neither ambiguous nor opaque;} \\ 6 & \text{if } f \text{ is } \mathrm{GL}_2(\mathbb{Z})\text{-equivalent to } x^2+xy+y^2; \\ 4 & \text{if } f \text{ is ambiguous and opaque;} \\ 2 & \text{otherwise.} \end{cases}$$

We note that if a quadratic form f is opaque, then its discriminant is positive; hence no positive definite binary quadratic form f is opaque.

When f is positive definite, then the number of elements in $\mathcal{V}_f(\mathbb{Z})$ of bounded height is finite and in fact lie in an ellipse. We shall enumerate these elements in Section 6.

3. Outer parametrization of binary quartic forms with $J = 0$

In [21], we obtained a new parametrization of binary quartic forms with small Galois groups. Let h be an integral binary quadratic form given as

$$h(x,y) = h_2x^2 + h_1xy + h_0y^2,$$

and analogous expressions for $u(x,y)$ and $v(x,y)$. Next put $\mathcal{J}(u,v)$ for the *Jacobian determinant* of u and v, given by

$$\mathcal{J}(u,v)(x,y) = \frac{1}{2}\begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = (u_2v_1 - u_1v_2)x^2 + 2(u_2v_0 - u_0v_2)xy + (u_1v_0 - u_0v_1)y^2. \tag{3.1}$$

We say that a pair of integral binary quadratic forms (u,v) is *primitive* if the gcd of the coefficients of $\mathcal{J}(u,v)$ is at most 2.

For a binary quartic form F, define its *cubic resolvent* to be the cubic polynomial

$$\mathcal{Q}_F(x) = x^3 - 3I(F)x + J(F). \tag{3.2}$$

We obtained the following in [21]:

Proposition 3.1 (Tsang, Xiao)**.** *Let F be a binary quartic with integer coefficients and non-zero discriminant. Then $\mathcal{Q}_F(x)$ is reducible over $\mathbb{Q}$ if and only if there exists an integral binary quadratic form h and a pair of primitive integral binary quadratic forms u, v such that*

$$F(x,y) = h(u(x,y), v(x,y)). \tag{3.3}$$

We further see that $J(F) = 0$ with $\mathcal{J}(u,v)^2|H_F$ if and only if

$$\Delta(v)h_0 - \Delta(u,v)h_1 + \Delta(u)h_2 = 0, \tag{3.4}$$

where

$$\Delta(u,v) = 2u_2v_0 - u_1v_1 + 2u_0v_2 \tag{3.5}$$

is the *joint discriminant* of the pair (u,v) of binary quadratic forms. We now give a brief overview of the invariant theory of *pairs* of binary quadratic forms.

3.1. **The action of $\mathrm{GL}_2(\mathbb{R}) \times \mathrm{GL}_2(\mathbb{R})$ on pairs of binary quadratic forms.** We shall denote by $\mathcal{U}_{2,2}(\mathbb{R})$ to be the six-dimensional $\mathbb{R}$-vector space of pairs of binary quadratic forms. That is,

$$\mathcal{U}_{2,2}(\mathbb{R}) = \left\{ \left(\begin{pmatrix} f_2 & f_1/2 \\ f_1/2 & f_0 \end{pmatrix}, \begin{pmatrix} g_2 & g_1/2 \\ g_1/2 & g_0 \end{pmatrix} \right) : f_2, f_1, f_0, g_2, g_1, g_0 \in \mathbb{R} \right\}. \tag{3.6}$$

The group $G(\mathbb{R}) = \mathrm{GL}_2(\mathbb{R}) \times \mathrm{GL}_2(\mathbb{R})$ acts on $\mathcal{U}_{2,2}(\mathbb{R})$ as follows. For $T = \left(\begin{smallmatrix} t_1 & t_2 \\ t_3 & t_4 \end{smallmatrix}\right)$ and $S \in \mathrm{GL}_2(\mathbb{R})$, with S^t denoting the transpose of S, we have

$$(T, S) \star (A, B) = \left(t_1 SAS^t + t_2 SBS^t, t_3 SAS^t + t_4 SBS^t \right).$$

The actions of the two copies of $\mathrm{GL}_2(\mathbb{R})$ commute, and we refer to the action of the first copy of $\mathrm{GL}_2(\mathbb{R})$ the *outer action* and the action of the second copy the *inner action*.

We now have the following:

Lemma 3.2. *Let $T \in \mathrm{GL}_2(\mathbb{R})$ be such that $\det T = \pm 1$. Put $(U, V) = (T, I_{2\times 2}) \star (u, v)$. Then $\mathcal{J}(U, V) = \mathcal{J}(u, v)$.*

Proof. Simple numerical verification. □

We next define the *invariant quadratic form* of a pair of quadratic forms (f, g), which is given as

$$\begin{aligned} \mathcal{F}(x, y) = \mathcal{F}_{(u,v)}(x, y) &= -\det \left(\begin{pmatrix} 2u_2 & u_1 \\ u_1 & 2u_0 \end{pmatrix} x - \begin{pmatrix} 2v_2 & v_1 \\ v_1 & 2v_0 \end{pmatrix} y \right) \\ &= \Delta(u)x + 2\Delta(u, v)xy + \Delta(v)y^2, \end{aligned}$$

We then have the following lemma:

Lemma 3.3. *The polynomials $\Delta(u), \Delta(u, v), \Delta(v)$ are the generators of the ring of polynomial invariants of the inner action of $\mathrm{GL}_2(\mathbb{R})$ on the set of pairs of binary quadratic forms. In particular, $\mathcal{F}(x, y)$ is invariant with respect to inner action.*

This result is classical; see for example [16]. See also [2] for a modern view.

A simple calculation reveals the following:

Lemma 3.4. *Let (u, v) be a pair of binary quadratic forms. Then*

$$\Delta(\mathcal{F}) = 4\Delta(\mathcal{J}(u, v)).$$

Furthermore, it is easy to check that the outer action on the pair (f, g) induces the usual substitution action of $\mathrm{GL}_2(\mathbb{R})$ on $\mathcal{F}(x, y)$.

Our goal is to show that when $f = \mathcal{J}(u, v)$ is positive definite (hence, the invariant quadratic form $\mathcal{F}(x, y)$ is necessarily positive definite by Lemma 3.4), there is essentially a *unique* choice of a pair of quadratic forms (u, v) such that (3.3) holds and both $\mathcal{J}(u, v)$ and $\mathcal{F}(x, y)$ are reduced. We shall prove the following:

Proposition 3.5. *Let F be a binary quartic form with integer coefficients, non-zero discriminant, and $J(F) = 0$. Then there exists a primitive pair of binary quadratic forms (u, v) such that $\mathcal{J}(u, v)$ and $\mathcal{F}_{(u,v)}$ are both reduced with positive leading coefficients and integers h_2, h_1, h_0 such that (3.3) holds. Moreover, the pair (u, v) is uniquely determined up to the action of $\mathrm{Aut}_{\mathbb{Z}}(\mathcal{F}_{(u,v)}) \times \mathrm{Aut}_{\mathbb{Z}}(\mathcal{J}(u, v)) \subset G(\mathbb{Z})$.*

Proof. Given any pair (u, v) for which (3.3) holds, the outer action of $\mathrm{GL}_2(\mathbb{Z})$ induces a change of variables of the quadratic form h; so any outer translate produces another representation of the shape (3.3). Similarly, inner action preserves the representability of F in the shape (3.3). Now suppose that both $\mathcal{J}(u, v)$ and $\mathcal{F}_{(u,v)}$ are fixed. Then the outer action is restricted to the subset of $\mathrm{GL}_2(\mathbb{Z})$ which fixes $\mathcal{F}_{(u,v)}$, in other words, $\mathrm{Aut}_{\mathbb{Z}}(\mathcal{F})$. Similarly, inner action is restricted to $\mathrm{Aut}_{\mathbb{Z}}(\mathcal{J}(u, v))$. □

We make the trivial observation that for any $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbb{Z})$, any binary quadratic form h and any pair of quadratic forms (u, v), we have

$$h(u, v) = h_{M^{-1}}(au + bv, cu + vd). \tag{3.7}$$

4. $\mathrm{SL}_2(\mathbb{Z})$-classes of binary quadratic forms and the Picard group of quadratic orders

Since $\mathcal{V}_f(\mathbb{Z})$ is canonically isomorphic to $\mathcal{V}_g(\mathbb{Z})$ whenever f and g are $\mathrm{GL}_2(\mathbb{Z})$-equivalent, it is thus pertinent to examine the properties of $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of binary quadratic forms. There is a rich history to this subject, and we will only pick from it what we need for the present work. See [2] for a modern treatment.

For technical reasons, we shall deal with $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes of binary quadratic forms. For a binary quadratic form f, denote by $[f]_{\mathbb{Z}}$ its $\mathrm{SL}_2(\mathbb{Z})$-equivalence class, and denote by $\mathcal{W}_2^*(\mathbb{Z})$ the set of $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes of binary quadratic forms. Put

$$\mathcal{W}_2^*(D) = \{w \in \mathcal{W}_2^*(\mathbb{Z}) : \Delta(w) = D\}$$

and put $\mathcal{O}_D$ for the unique quadratic order of discriminant D. It is well-known that the set of primitive classes $[f]_{\mathbb{Z}}$ with discriminant D parametrize the ideal classes in the Picard group $\mathrm{Pic}(\mathcal{O}_D)$, the group of ideal classes in $\mathcal{O}_D$ (see [2] for a modern treatment). We put

$$h_2(D) = \#\,\mathrm{Pic}(\mathcal{O}_D). \tag{4.1}$$

We then have the following famous theorem, originally conjectued by Gauss in [13] and subsequently proved by Mertens and Siegel [18]:

Proposition 4.1 (Gauss, Mertens, Siegel)**.** *The class number $h_2(-D)$ satisfies the following asymptotic formulas:*

$$\sum_{0<D\leq X} h_2(-D) = \frac{\pi}{18\zeta(3)}X^{3/2} + O(X\log X) \tag{4.2}$$

and

$$\sum_{\substack{0<D\leq X\\ D\equiv 0 \pmod 4}} h_2(-D) = \frac{\pi}{42\zeta(3)}X^{3/2} + O(X\log X). \tag{4.3}$$

For $D>0$, we put $R_D = \log\varepsilon_D$ for the regulator *of the quadratic field $\mathbb{Q}(\sqrt{D})$. We then have*

$$\sum_{0<D\leq X} h_2(D)R_D = \frac{\pi^2}{18\zeta(3)}X^{3/2} + O\left(X\log X\right) \tag{4.4}$$

and

$$\sum_{\substack{0<D\leq X\\ D\equiv 0 \pmod 4}} h_2(D)R_D = \frac{\pi^2}{42\zeta(3)}X^{3/2} + O\left(X\log X\right). \tag{4.5}$$

We remark that our class number (4.1) only counts *primitive* classes.

Recall that a positive definite binary quadratic form $f(x,y) = \alpha x^2 + \beta xy + \gamma y^2$ is said to be *reduced* if its coefficients satisfy $|\beta| \leq \alpha \leq \gamma$. Gauss proved that $h_2(-D)$ is exactly equal to the number of primitive reduced forms of discriminant $-D$.

We now require the following lemma, which is useful when estimating the sum of the error terms as we sum across $\mathrm{SL}_2(\mathbb{Z})$-classes over positive definite binary quadratic forms:

Lemma 4.2. *Let $g(x,y) = g_2x^2 + g_1xy + g_0y^2$ be a positive definite and reduced binary quadratic form with co-prime integer coefficients. Let $\sum_{D\leq Y}^{\dagger}$ denote the sum over positive definite, reduced, and primitive binary quadratic forms g of discriminant up to Y. Then*

$$\sum_{D\leq X^{2/3}}^{\dagger} \frac{1}{(g_2D)^{1/2}} = O\left(X^{1/2}\right). \tag{4.6}$$

Proof. We note that the sum (4.6) is approximated by the integral

$$\mathfrak{I}(X) = \int_1^{X^{1/3}} \int_a^{X^{2/3}/a} \int_{-a}^{a} \frac{dbdcda}{a\sqrt{c}}.$$

We evalute $\mathfrak{I}(X)$ by

$$\begin{aligned}\mathfrak{I}(X) &= \int_1^{X^{1/3}} \int_a^{X^{2/3}/a} \frac{2dcda}{\sqrt{c}} \\ &= \int_1^{X^{1/3}} 4\left(\frac{X^{1/3}}{\sqrt{a}} - a^{1/2}\right) da \\ &= O\left(X^{1/3} \cdot X^{1/6}\right) = O\left(X^{1/2}\right),\end{aligned}$$

as desired. □

Next we state a similar result to Proposition 4.1 for $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes of reducible forms.

Proposition 4.3. *Let n be a positive integer. The number of $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes of primitive, integral binary quadratic forms of discriminant n^2 is equal to $\phi(n)$, and an explicit set of representatives is*

$$\{ax^2 + nxy : 1 \leq a \leq n-1, \gcd(a,n) = 1\}.$$

Therefore,

$$\sum_{n \leq X^{1/2}} h_2(n^2) = \sum_{n \leq X^{1/2}} \phi(n) = \frac{3X}{\pi^2} + O\left(X^{1/2} \log X\right). \tag{4.7}$$

Finally, we have a similar proposition for forms $\mathrm{GL}_2(\mathbb{Q})$-equivalent to the form $x^2 + y^2$:

Proposition 4.4. *Let n be a positive integer. The number of $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes fo primitive, integral binary quadratic forms of discriminant $-4n^2$ is given by*

$$h_2(-4n^2) = n \prod_{p \equiv 1,2 \pmod 4} \left(1 - \frac{1}{p}\right) \prod_{p \equiv 3 \pmod 4} \left(1 + \frac{1}{p}\right).$$

We further have the asymptotic formula

$$\sum_{n \leq X^{1/2}/2} h_2(-4n^2) = \frac{3}{32}\left(\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^2}\right)^{-1} X + O\left(X^{1/2} \log X\right). \tag{4.8}$$

Proof. The first statement is found in [8], pages 109-119. We now prove the asymptotic formula.

We denote by $\rho(n) = h_2(-4n^2)$. Then $\rho(n)$ is a multiplicative function and the Dirichlet series of $\rho(n)$ converges absolutely and is holomorphic for $s > 2$. It admits a factorization into the Euler product

$$\zeta(s-1) \prod_{p \equiv 1,2 \pmod 4} (1 - p^{-s}) \prod_{p \equiv 3 \pmod 4} (1 + p^{-s}). \tag{4.9}$$

This follows from the identities

$$\sum_{n=1}^{\infty} \frac{\phi(n)}{n^s} = \frac{\zeta(s-1)}{\zeta(s)} \text{ and } \sum_{n=1}^{\infty} \frac{\psi(n)}{n^s} = \frac{\zeta(s)\zeta(s-1)}{\zeta(2s)},$$

where $\psi(n) = n\prod_{p|n}(1 + 1/p)$. Let χ be the unique non-principal character of modulus 4. Then (4.9) can be written as

$$\zeta(s-1)L(s,\chi)^{-1}(1 - 2^{-s}) = \zeta(s-1)\beta(s)^{-1}(1 - 2^{-s}), \tag{4.10}$$

where $\beta(s)$ is Dirichlet's beta series. The asymptotic formula (4.8) then follows from Perron's formula. □

We will use the results in this section to allow us to sum over different error terms that arise in the proof of Theorem 1.1.

5. Some arithmetic and algebraic properties of binary quadratic forms

We will be counting with respect to the I-invariant. We now give the I-invariant of F for F given as in Proposition 2.3. By (1.6), for all $F \in \mathcal{V}_4(\mathbb{R})$ we have

$$\Delta(F) = \frac{4I(F)^3}{27},$$

since $J(F) = 0$. Therefore the condition $|\Delta(F)| \leq X$ is translated into

$$|I(F)|^3 \leq \frac{27X}{4}. \tag{5.1}$$

We note that for $\Delta(F) > 0$ for all $F \in \mathcal{V}_4(\mathbb{R})$, since the real quadratic form divisor of the Hessian is positive definite. We may thus drop the absolute value in (5.1). We have the following lemma:

Lemma 5.1. *Let $f(x, y) = \alpha x^2 + \beta xy + \gamma y^2$ be a binary quadratic form with $\alpha\Delta(f) \neq 0$. Then for F given as in Proposition 2.3, we have*

$$I(F) = \frac{-3(\alpha B^2 - 4\beta AB + 16\gamma A^2)\Delta(f)}{4\alpha^3}. \tag{5.2}$$

We now take

$$\mathcal{I}(F) = \frac{\alpha B^2 - 4\beta AB + 16\gamma A^2}{4\alpha^3}. \tag{5.3}$$

Now put

$$N_f(X) = \#\{F \in \mathcal{V}_f(\mathbb{Z}) : \mathcal{I}(F) \leq X\}. \tag{5.4}$$

Since $N_f(X) = N_g(X)$ whenever f, g are $\mathrm{GL}_2(\mathbb{Z})$-equivalent, we shall assume from now on that f takes on a convenient form. We shall now take a $\mathrm{GL}_2(\mathbb{Z})$-translate of f to be

$$px^2 + mxy + ny^2, \tag{5.5}$$

Here p is the smallest odd prime representable by f not dividing D and n is the smallest positive integer for which (5.5) holds, and then we choose m to be non-negative. This translate of f is then determined given p.

We now show, at least when f is given in (5.5), that whenever $F \in \mathcal{V}_f(\mathbb{Z})$ that $\mathcal{I}(F) \in \mathbb{Z}$. We require the following result, which is a special case of Theorem 2 in [19]:

Lemma 5.2. *Let $f = \alpha x^2 + \beta xy + \gamma y^2$ be a primitive binary quadratic form with integer coefficients and non-zero discriminant d, and let p be an odd prime which is representable by a quadratic form of discriminant d which does not divide d. Then there exist linear forms L_1, L_2 with coefficients in the p-adic integers $\mathbb{Z}_p$ such that*

$$f(x, y) = L_1(x, y)L_2(x, y)$$

over $\mathbb{Z}_p$. Further, for any positive integer k, the solutions to the congruence $f(x, y) \equiv 0 \pmod{p^k}$ with x, y not both zero modulo p lie in exactly one of the two lattices

$$\mathcal{L}_1^k = \{(x, y) \in \mathbb{Z}^2 : L_1(x, y) \equiv 0 \pmod{p^k}\}$$

and

$$\mathcal{L}_2^k = \{(x, y) \in \mathbb{Z}^2 : L_2(x, y) \equiv 0 \pmod{p^k}\}.$$

.

We call the lattice $\mathcal{L}_i^k$ the k-th *Hensel lift* of the lattice $\mathcal{L}_i$. We now prove that the $I(F)/\Delta(f) = \mathcal{I}(F)$ is always an integer whenever $F \in \mathcal{V}_f(\mathbb{Z})$.

Lemma 5.3. *Let f be given as in (5.5), and let $\mathcal{I}(F)$ be given as in (5.3). Then $\mathcal{I}(F) \in \mathbb{Z}$ whenever $F \in \mathcal{V}_f(\mathbb{Z})$.*

Proof. With f in the form given in (5.5), the congruence condition (2.4) implies (2.3) and (2.2). The only two primes that need to be considered are 2 and p itself. Note that since we assumed $p \nmid \Delta(f)$, it follows that $p \nmid m$. Therefore examining $\mathcal{A}_3 \equiv 0 \pmod{p}$ yields

$$\begin{aligned} 4n(m^2 - pn)A - m(m^2 - 2pn)B &\equiv m^2(4nA - mB) \\ &\equiv 0 \pmod{p}, \end{aligned}$$

hence

$$4nA - mB \equiv 0 \pmod{p}$$

or (2.2). Now reducing modulo p^2 we get

$$(4m^2nA - m^3B + mnpB) - pn(4nA - mB) \equiv 0 \pmod{p^2},$$

and the second term vanishes mod p^2, hence the first term must vanish as well. However the first term is equal to $m(4mnA - (m^2 - pn)B)$, which implies (2.3). Moreover, we see that (2.4) is soluble with $A \not\equiv 0 \pmod{p}$, whence $\mathcal{L}_{f,p}$ is contained in the 3rd Hensel lift of the lattice defined by (2.2), whence $p^3 | pB^2 - 4mAB + 16nB^2$ whenever $(A, B) \in \mathcal{L}_{f,p}$.

We now have to deal with the prime 2. If m is odd, then $\mathcal{A}_3 \equiv 0 \pmod{4}$ is equivalent to $B \equiv 0 \pmod{4}$ and $\mathcal{A}_1 \equiv 0 \pmod{2}$ is equivalent to $B \equiv 0 \pmod{2}$. We then see at once this is sufficient for the numerator of $\mathcal{I}(F)$ to be divisible by 4. This proves the claim. □

We can now clear the denominator in (5.3) by restricting to the lattice $\mathcal{L}_{f,p}$ to get a new quadratic form $\nu(f)$. A potential problem is that we do not know about the $\mathrm{SL}_2(\mathbb{Z})$-equivalence class of $\nu(f)$ given f. It turns out that this would create problems when estimating the contribution of $N_f(X)$ with $D = -\Delta(f)$ large. Therefore, we must resolve this issue.

5.1. **Auxiliary quadratic forms $w(f)$ and $\nu(f)$.** We now define certain auxiliary quadratic forms $w(f), \nu(f)$ for quadratic forms given in (5.5). Suppose that m is odd. Then for a quartic form F given by (2.6) to have integer coefficients, we must have $B \equiv 0 \pmod{4}$. No congruence conditions modulo a power of 2 is imposed if m is even. Thus, when m is odd we shall assume $B \equiv 0 \pmod{4}$, which changes (5.2) to

$$\frac{4(pB^2 - mAB + nA^2)\Delta(f)}{p^3}. \tag{5.6}$$

We then put

$$w(f)(x, y) = \begin{cases} px^2 - mxy + ny^2 & \text{if } m \text{ is odd} \\ px^2 - 4mxy + 16ny^2 & \text{if } m \text{ is even.} \end{cases} \tag{5.7}$$

When m is odd, put

$$\mathcal{L}_2(w(f)) = \{(x, y) \in \mathbb{Z}^2 : mx \equiv ny \pmod{p}\}, \tag{5.8}$$

and $\mathcal{L}_2^3(w(f))$ for its 3rd Hensel lift. Likewise, when m is even, put

$$\mathcal{L}_2(w(f)) = \{(x, y) \in \mathbb{Z}^2 : mx \equiv 4ny \pmod{p}\} \tag{5.9}$$

and $\mathcal{L}_2^3(w(f))$ for its 3rd Hensel lift. If $\left\{ \begin{pmatrix} u_1 \\ v_1 \end{pmatrix}, \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} \right\}$ is a basis for $\mathcal{L}_2^3(w(f))$, then the matrix $U = \left(\begin{smallmatrix} u_1 & u_2 \\ v_1 & v_2 \end{smallmatrix} \right)$ satisfies

$$w(f)(u_1x + u_2y, v_1x + v_2y) = p^3 g(x, y) \tag{5.10}$$

for some primitive binary quadratic form g. We then have the following equality:

$$\{w(f)(x, y) : (x, y) \in \mathcal{L}_2^3(f)\} = \{p^3 g(u, v) : (u, v) \in \mathbb{Z}^2\}. \tag{5.11}$$

Moreover, g is well-defined up to $\mathrm{GL}_2(\mathbb{Z})$-equivalence and $\Delta(w(f)) = \Delta(g)$ by (51) in [19]. We shall denote this g (or any $\mathrm{GL}_2(\mathbb{Z})$-translate of it) by $\nu(f)$.

We may now state the main proposition of this subsection:

Proposition 5.4. *Let f be a positive definite binary quadratic form with co-prime integer coefficients and discriminant $-D$, and let $w(f)$ be given as in (5.7). Then there exists a quadratic form $\nu(f)$ which satisfies (5.11) with $\Delta(w(f)) = \Delta(\nu(f))$ and $[w(f)]_{\mathbb{Z}}^4 = [\nu(f)]_{\mathbb{Z}}$. Moreover, for primitive $g \in V_2(\mathbb{Z})$ with discriminant $\Delta(f)$, we have $w(f), w(g)$ are $\mathrm{GL}_2(\mathbb{Z})$-equivalent if and only if f, g are $\mathrm{GL}_2(\mathbb{Z})$-equivalent.*

The proof of Proposition 5.4 relies on a refinement of Hensel's lemma applied to prime ideals. We first recall the following classical fact:

Lemma 5.5. *Let f be a positive definite binary quadratic form with co-prime integer coefficients and discriminant $-D$, and let m be a positive integer. Then m can be represented by f if and only if the principal ideal (m) admits a factorization as $\mathfrak{m}\overline{\mathfrak{m}}$, where $\mathfrak{m}$ is an ideal in $\mathcal{O}_{-D}$ in the ideal class parametrized by $[f]_{\mathbb{Z}}$.*

Note that for a prime p, the principal ideal (p) is either inert in $\mathcal{O}_{-D}$ or splits into $\mathfrak{p}_1\mathfrak{p}_2$. Then p is representable by f if and only if it splits in $\mathcal{O}_{-D}$ and either $\mathfrak{p}_1$ or $\mathfrak{p}_2$ lies in the ideal class parametrized by $[f]_{\mathbb{Z}}$. Note that the ideal class of $\mathfrak{p}_1$ is the inverse of $\mathfrak{p}_2$ in the ideal class group, since their product is a principal ideal. We thus have the following corollary to Lemma 5.5:

Corollary 5.6. *Le D be a positive integer and let p be a prime not dividing D. Then p is represented by exactly two $\mathrm{SL}_2(\mathbb{Z})$-classes of primitive quadratic forms of discriminant $-D$ up to multiplicity, and these ideal classes are inverses of each other in the ideal class group.*

Now note the following simple observation: if $f(x,y) = ax^2+bxy+cy^2$, then a representative of its inverse class $[f]_{\mathbb{Z}}^{-1}$ is given by $ax^2 - bxy + cy^2$. In particular, the classes $[f]_{\mathbb{Z}}$ and $[f]_{\mathbb{Z}}^{-1}$ are $\mathrm{GL}_2(\mathbb{Z})$-equivalent.

We now show that the forms $w(f)$ given in (5.7) are distinct, which proves one part of Proposition 5.4.

Lemma 5.7. *Let f, g be two binary quadratic forms with co-prime integer coefficients and equal discriminant. Then the forms $w(f), w(g)$ given in (5.7) are $\mathrm{GL}_2(\mathbb{Z})$-distinct if and only if f and g are $\mathrm{GL}_2(\mathbb{Z})$-distinct.*

Proof. It is clear that if f and g are $\mathrm{GL}_2(\mathbb{Z})$-equivalent, then so are $w(f)$ and $w(g)$. For the converse, first suppose that m is odd. Then $w(f)$ is $\mathrm{GL}_2(\mathbb{Z})$-equivalent to f via $\left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$, hence the statement is clear. Now suppose that m is even. Put $g = g_2x^2 + g_1xy + g_0y^2$, where g_2 is an odd prime. Note that g_1 is even, since $\Delta(f) = \Delta(g)$ and the parity only depends on the middle coefficient. It then follows that

$$w(g) = g_2x^2 - 4g_1xy + 16g_0y^2.$$

We suppose that $w(f)$ is equivalent to $w(g)$, and let $T = \left(\begin{smallmatrix}t_1 & t_2\\ t_3 & t_4\end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbb{Z})$ be such that

$$w(f)_T(x,y) = p(t_1x+t_2y)^2 - 4m(t_1x+t_2y)(t_3x+t_4y) + 16n(t_3x+t_4y)^2 = w(g)(x,y).$$

Since g_2 is odd, it follows that t_1 is odd. Moreover, the middle coefficient of $w(f)_T$ is equal to

$$2(pt_1t_2 - 2mt_2t_3 - 2mt_1t_4 + 16nt_3t_4).$$

We need this to be divisible by 8, since $8|4g_1$. This implies that $2pt_1t_2 \equiv 0 \pmod 8$. However p, t_1 are odd, so $t_2 \equiv 0 \pmod 4$. It then follows that

$$T' = \begin{pmatrix}1 & 0\\ 0 & 4\end{pmatrix} T \begin{pmatrix}1 & 0\\ 0 & 1/4\end{pmatrix} = \begin{pmatrix}t_1 & t_2/4\\ 4t_3 & t_4\end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z}),$$

which shows that f and g are equivalent. □

5.2. **Hensel lifting of lattices and $\mathrm{SL}_2(\mathbb{Z})$-classes of binary quadratic forms.** We shall treat forms f given in the shape (5.5) (note that $w(f)$ is also of the shape (5.5)). Put

$$\Lambda_1(f) = \{(x,y) \in \mathbb{Z}^2 : y \equiv 0 \pmod p\} \tag{5.12}$$

and

$$\Lambda_2(f) = \{(x,y) \in \mathbb{Z}^2 : mx + ny \equiv 0 \pmod p\}. \tag{5.13}$$

Following the notation of Lemma 5.2, we put $\Lambda_i^k(f)$ for the k-th Hensel lift of the lattice $\Lambda_i(f)$.

For each i and k, we assign a class $w \in W_2(\mathbb{Z})$ to Λ_i^k as follows. There exists a quadratic form $g_{i,k}$, unique up to $\mathrm{GL}_2(\mathbb{Z})$-equivalence, such that

$$\{f(x,y) : (x,y) \in \Lambda_i^k\} = \{p^k g_{i,k}(x,y) : (x,y) \in \mathbb{Z}^2\}. \tag{5.14}$$

The form $g_{i,k}$ has the same discriminant as f and hence $[g_{i,k}]_{\mathbb{Z}}$ is in the same ideal class group as $[f]_{\mathbb{Z}}$.

Lemma 5.8. *Let f be given as in (5.5). Then for each $k \geq 1$ one can choose an integral binary quadratic form $g_{i,k}$ satisfying (5.14) such that $[g_{1,k}]_{\mathbb{Z}} = [f]_{\mathbb{Z}}^{k-1}$ and $[g_{2,k}]_{\mathbb{Z}} = [f]_{\mathbb{Z}}^{k+1}$.*

We shall state a slightly more general result, which may be of separate interest. From here on, f shall not be assumed to take the form (5.5). We hence return to the notation (2.1) for f.

Proposition 5.9. *Let f be a binary quadratic form with co-prime integer coefficients and non-zero discriminant. Suppose that p is an odd prime such that $\left(\frac{\Delta(f)}{p}\right) = 1$, and let $\mathfrak{p}_1, \mathfrak{p}_2$ be the two prime ideal divisors of (p). Suppose further that there exists a non-negative integer s for which $[f]_{\mathbb{Z}} = [\mathfrak{p}_1^s]$. Let $\mathcal{L}_1, \mathcal{L}_2$ be as given in Lemma 5.2. For all $k \geq 1$, there exists integral quadratic forms $g_{1,k}, g_{2,k}$ such that*

$$\{f(x,y) : (x,y) \in \mathcal{L}_i^k\} = \{p^k g_{i,k}(x,y) : (x,y) \in \mathbb{Z}^2\}$$

for $i = 1, 2$ and

$$[g_{1,k}]_{\mathbb{Z}} = [\mathfrak{p}_1]^{s-k}, [g_{2,k}]_{\mathbb{Z}} = [\mathfrak{p}_1]^{s+k}.$$

Proof. See the Appendix by Erick Knight. □

We then have the following corollary, which a crucial component of our proof:

Corollary 5.10. *Let f be given as in (5.5), $w(f)$ as given in (5.7). Then $\nu(f)$ can be chosen so that $[\nu(f)]_{\mathbb{Z}} = [w(f)]_{\mathbb{Z}}^4$.*

The proof of Proposition 5.4 then follows from Corollary 5.10 and Lemma 5.7.

6. Enumerating the elements in $\mathcal{V}_f(\mathbb{Z})$ for a fixed positive definite quadratic form f

In this section we shall give an asymptotic formula for $N_f(X)$.

Theorem 6.1. *Let $f(x,y) = \alpha x^2 + \beta xy + \gamma y^2$ be a positive definite reduced binary quadratic form with co-prime integer coefficients. Put $D = |\Delta(f)|$. Then*

$$N_f(X/D) = \begin{cases} \dfrac{\pi X}{3D^{3/2}} + O\left(\dfrac{X^{1/2}}{D^{1/2}}\right) & \text{if } \beta \text{ is odd;} \\ \\ \dfrac{4\pi X}{3D^{3/2}} + O\left(\dfrac{X^{1/2}}{D^{1/2}}\right) & \text{if } \beta \text{ is even.} \end{cases} \tag{6.1}$$

The following lemma illustrates the importance of having f as a positive definite form.

Lemma 6.2. *Let f be a primitive, positive definite binary quadratic form with integer coefficients. Then the I-invariant of F is a positive definite binary quadratic form and hence the set of $F \in \mathcal{V}_f(\mathbb{R})$ with $I(F) \leq X^{1/3}$ lie in the ellipse defined by*

$$\mathcal{E}_f(X) = \left\{(A,B) \in \mathbb{R}^2 : \alpha B^2 - 4\beta AB + 16\gamma A^2 \leq \frac{4\alpha^3}{3|\Delta(f)|} X^{1/3}\right\}. \tag{6.2}$$

Proof. The fact that the I-invariant is a positive definite binary quadratic form follows from (5.2), namely the observation that the I-invariant is a positive multiple of the quadratic form $\alpha x^2 - 4\beta xy + 16\gamma y^2$, which is equal to $f(x, -4y)$; hence positive definite. □

Since $N_f(X)$ only depends on the class $\mathcal{C}(f)$ of f, we may pick a convenient representative of f. Indeed, we may suppose that α is odd and co-prime to $\Delta(f)$, which implies that $\gcd(\alpha, \beta) = 1$.

It follows from Lemma 6.2 and Proposition 2.3 that

$$\{F \in \mathcal{V}_{f,\mathbb{Z}} : I(F) \leq X\} = \mathcal{E}_f \cap \mathcal{L}_{f,\alpha}.$$

We then need to compute the determinant $\det(\mathcal{L}_{f,\alpha})$, which we do so in the following proposition.

Proposition 6.3. *The determinant $\det \mathcal{L}_{f,\alpha}$ is equal to $4\alpha^3$ if β is odd and α^3 otherwise.*

Proof. Let k be the exponent of p dividing α. For $p \geq 3$, the congruence (2.4) implies

$$\beta(4\beta\gamma A - (\beta^2 - \alpha\gamma)B) - \alpha\gamma(4\gamma A - \beta B) \equiv 0 \pmod{p^{3k}}.$$

It follows that

$$4\beta\gamma A - (\beta^2 - \alpha\gamma)B \equiv 0 \pmod{p^k},$$

which is equivalent to

$$4\gamma A - \beta B \equiv 0 \pmod{p^k},$$

or $\mathcal{A}_1 \equiv 0 \pmod{p^k}$. This then implies that

$$\beta(4\beta\gamma A - (\beta^2 - \alpha\gamma)B) \equiv 0 \pmod{p^{2k}},$$

or $\mathcal{A}_2 \equiv 0 \pmod{p^{2k}}$. Thus (2.4) implies (2.3) and (2.2), so that

$$\det \mathcal{L}_{f,\alpha}^{(p)} = p^{3k}.$$

We now deal with the case when $p = 2$. If $2 \nmid \beta$ then the argument proceeds as before, but the congruences modulo p^k, p^{2k}, p^{3k} are replaced by $2^{k+1}, 2^{2k}, 2^{3k+2}$ respectively. □

We may now prove Theorem 6.1.

Proof of Theorem 6.1. We wish to count the number of points in the intersection $\mathcal{E}_f(X) \cap \mathcal{L}_{f,\alpha}$. We apply Davenport's lemma, which asserts that

$$\#\mathcal{E}_f(X) \cap \mathcal{L}_{f,\alpha} = \frac{\mathrm{Vol}(\mathcal{E}_f(X))}{|\det(\mathcal{L}_{f,\alpha})|} + O\left(\max\left\{\mathrm{Vol}(\overline{\mathcal{E}_f(X)}), 1\right\}\right).$$

The volume of $\mathcal{E}_f(X)$ is given by

$$\mathrm{Vol}(\mathcal{E}_f(X)) = \frac{4\alpha^3}{3|\Delta(f)|^{3/2}} X^{1/3},$$

whence

$$N_f^{(0)}(X) = \begin{cases} \dfrac{\pi X^{1/3}}{3|\Delta(f)|^{3/2}} + O\left(\dfrac{X^{1/6}}{|\Delta(f)|^{1/2}}\right) & \text{if } \beta \text{ is odd;} \\ \dfrac{4\pi X^{1/3}}{3|\Delta(f)|^{3/2}} + O\left(\dfrac{X^{1/6}}{|\Delta(f)|^{1/2}}\right) & \text{if } \beta \text{ is even,} \end{cases}$$

since $\alpha \leq \gamma$ since f is reduced.

□

The task now, given Theorem 6.1, is to obtain uniformity estimates for the error term that appears. We will need the following lemma, essentially Lemma 3.1 in [6].

Lemma 6.4. *Let $f(x,y) = \alpha x^2 + \beta xy + \gamma y^2$ be a positive definite integral binary quadratic form which is reduced. Then the set of integer pairs (x,y) satisfying $f(x,y) \leq X$ is given by*

$$\frac{2\pi X}{\sqrt{4\alpha\gamma - \beta^2}} + O\left(\sqrt{\frac{X}{\alpha}}\right).$$

7. Two ways to count quartic forms with vanishing J-invariant

In Section 6 we showed how to estimate the quantities $N_f(X)$. By summing over $\mathrm{SL}_2(\mathbb{Z})$-classes of f having discriminant $-D$ bounded by X, we will be able to prove Theorem 1.1. However, this is infeasible: as soon as $D > X^{2/3}$ the main term in (6.1) will be less than one. It then becomes un clear how continuing to sum $N_f(X)$ past that point will contribute to a main term.

We put $N(X;Y)$ for the quantity (1.8), with the additional stipulation that we are only counting contributions from those $N_f(X)$ with $|\Delta(f)| \leq Y$. Using the theory introduced in Section 3, we will introduce a different way to enumerate elements in $\mathcal{W}_4^{(0)}(\mathbb{Z})(X)$, via the *outer form* $h(x,y)$ in (3.3). In particular, put

(7.1)
$$\mathcal{S}_h(X) = \#\{w \in \mathcal{W}_4^{(0)}(\mathbb{Z}) : 0 < |\Delta(w)| \leq X, \exists \text{ primitive quadratic forms } u, v \text{ s.t. } h(u(x,y), v(x,y)) \in w\}$$

Recall (3.4) the relation

$$\Delta(u)h_2 - \Delta(u,v)h_1 + \Delta(v)h_0 = 0.$$

Using the fact that $-D = \Delta(\mathcal{J}(u,v))/4$, we see that we can express $-D$ as a function of the quadratic form $h(x,y) = h_2x^2 + h_1xy + h_0y^2$ and $\Delta(u), \Delta(u,v), \Delta(v)$, namely by setting

$$\Delta(v) = -\frac{h_2\Delta(u) - h_1\Delta(u,v)}{h_0}.$$

This gives

$$-D = \frac{h_2\Delta(u)^2 - h_1\Delta(u)\Delta(u,v) + h_0\Delta(u,v)^2}{h_0} = \frac{h(\Delta(u), -\Delta(u,v))}{h_0}. \tag{7.2}$$

Thus, we have the key equation

$$I(F) = \frac{-3h(\Delta(u), -\Delta(u,v))}{h_0} \cdot \Delta(h). \tag{7.3}$$

Thus we can formulate the question as finding pairs $(x, y) \in \mathbb{Z}^2$ for which

$$h(x,y) \equiv 0 \pmod{h_0} \tag{7.4}$$

and such that

$$\left|\frac{h(x,y)}{h_0}\right| \leq \frac{X}{\Delta(h)}. \tag{7.5}$$

The caveat is that not all such pairs (x, y) are admissible: indeed, we shall only take those pairs (x, y) satisfying the congruence condition (7.4) and such that one can find a primitive pair of quadratic forms (u, v) such that $\Delta(u) = x, \Delta(u, v) = -y$. Luckily such a criterion is already known, due to Morales [16]: such a pair exists if and only if x is a square modulo $-D$. We state Morales' result for convenience:

Proposition 7.1 (Morales [16])**.** *The number of inner equivalence classes of pairs (u, v) of integral binary quadratic forms with prescribed invariant form $\mathcal{F}(x, y) = \delta_1 x^2 + 2\delta_{1,2}xy + \delta_2 y^2$ is equal to*

$$\sum_{\substack{c \mid \delta_{1,2}^2 - \delta_1\delta_2 \\ c>0, c \text{ square-free}}} \left(\frac{\delta_1}{c}\right). \tag{7.6}$$

Note that since h is indefinite, the number of solutions to (7.5) is a priori infinite. Thus to make sense of the counting problem we must account for the action of the unit group of $\mathcal{O}_h = \mathcal{O}_{\mathbb{Q}(\sqrt{\Delta(h)})}$ on h, so that at most a bounded number of points from each orbit is counted. To do so we must define such an action, but unfortunately both the inner and outer actions introduced in Section 3 do not have an immediate interpretation in terms of the expression introduced in (7.2). Fortunately, with respect to the unit group action, the outer action induces the correct action on (7.2). Therefore, Proposition 7.1 implies the following:

Proposition 7.2. *Let $h(x, y)$ be a primitive integral binary quadratic form such that $\Delta(h) > 0$. Then*

$$\mathcal{S}_h(X) = \sum_{\substack{-Xh_0 < h(x,y) < 0 \\ h_2 x - h_1 y \equiv 0 \pmod{h_0}}}^{*} \sum_{\substack{c \mid h(x,y) \\ c>0, c \text{ square-free}}} \left(\frac{x}{c}\right), \tag{7.7}$$

where the summation $\sum^*$ denotes summing over a suitable fundamental domain $\mathfrak{D}$ for the action of the unit group in $\mathbb{Q}(\sqrt{\Delta(h)})$ on h.

What will be important for us is the following. Put $\mathcal{N}(X; Y)$ for the sum of $\mathcal{S}_h(X)$ over $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes of h with $E = \Delta(h) \leq Y$. We then have:

Proposition 7.3. *For any positive number Y we have*

$$N(X) = N(X; Y) + \mathcal{N}\left(X; XY^{-1}\right).$$

Proof. If $w \in \mathcal{W}_4^{(0)}(\mathbb{Z})$ is such that w is counted by both $N_f(X)$ and $\mathcal{S}_h(X)$, then

$$I(F) = -3\Delta(f)\Delta(h).$$

Thus, if $|I(F)| \leq X$, then F is counted by either $N_f(X)$ with $|\Delta(f)| \leq Y$ or $\mathcal{S}_h(X)$ with $|\Delta(h)| \leq XY^{-1}$. This completes the proof. □

Therefore to complete the proof of Theorem 1.1, we need to estimate $\mathcal{N}(X; X^{1/3})$. To do so, we need to estimate $\mathcal{S}_h(X)$ with reasonable precision.

Indeed a suitable choice of fundamental domain is crucial for the necessary estimates: for this we follow a construction due to Schmidt [17]. Recall that for $E = \Delta(h)$, we denote by R_E the regulator of the quadratic order $\mathcal{O}_E$ of discriminant E. Following [17], we put

$$t_E = \lfloor R_E \rfloor + 1, u = \exp(R_E/t_E).$$

Then we have

$$u^{t_E} = \varepsilon_E \text{ and } 1 \ll \log u \leq 1,$$

where ε_E is the fundamental unit of $\mathcal{O}_E$. Instead of considering sublattices of $\mathbb{Z}^2$, we instead consider, for points $\alpha \in K_E = \mathbb{Q}(\sqrt{E})$, the put

$$\widehat{\alpha} = (\alpha, \overline{\alpha}) \in \mathbb{R}^2,$$

where $\overline{\alpha}$ is the conjugate of α in $\mathcal{O}_E$. Note that the set of $\widehat{\alpha}, \alpha \in \mathcal{O}_E$ gives a sub-lattice of $\mathbb{R}^2$ of discriminant $E^{1/2}$. When α is restricted to a non-zero ideal $\mathfrak{a} \subset \mathcal{O}_E$, then $\widehat{\alpha}$ runs over a lattice $\Lambda(\mathfrak{a})$ satisfying $\det \Lambda(\mathfrak{a}) = E^{1/2}\mathfrak{N}(\mathfrak{a})$, where $\mathfrak{N}(\mathfrak{a})$ refers to the norm of the ideal $\mathfrak{a}$. Further, for $\mathfrak{C}$ an ideal class in $\mathcal{O}_E$ and $\mathfrak{c}_1, \mathfrak{c}_2, \cdots$ the integral ideals in $\mathfrak{C}$ ordered by norm, put

$$\mathfrak{N}(\mathfrak{C}) = \left(\sum_{j=1}^{2t} \mathfrak{N}(\mathfrak{c}_j) \right)^{-1/2}.$$

Put $v = u^{1/2}$, so that $1 \ll \log v$, and $v - 1 \gg 1$. Put

$$\tau : \mathbb{R}^2 \to \mathbb{R}^2, \tau(\alpha, \overline{\alpha}) = \left(v^{-1}\alpha, v\overline{\alpha}\right).$$

Put $\Lambda(\mathfrak{a}, j)$ for the image of $\Lambda(\mathfrak{a})$ under the map τ^j, where the exponent refers to functional composition. Then $\Lambda(\mathfrak{a}, j)$ is again a lattice in $\mathbb{R}^2$, since τ is a linear map. Moreover we have $\det \Lambda(\mathfrak{a}, j) = \det \Lambda(\mathfrak{a})$.

What we gain is that Schmidt shows in [17] that we have a very nice expression for the first successive minimum of $\Lambda(\mathfrak{a}, j)$, given by

$$\lambda_1(\mathfrak{a}, j) = \min_{\alpha \in \mathfrak{a} \setminus \{0\}} \left(v^{-2j}|\alpha|^2 + v^{2j}|\overline{\alpha}|^2\right)^{1/2}. \tag{7.8}$$

Now put, for $\alpha \in K_E$,

$$\psi(\alpha) = \frac{|\alpha|}{|\overline{\alpha}|}.$$

By explicit calculation we see that $\psi(\varepsilon_E \cdot \alpha) = \varepsilon_E^2 \psi(\alpha)$ for all $\alpha \in \mathcal{O}_E$. This shows that for each such α there exists uniquely an integer s such that

$$\varepsilon_E^{-1} < \psi(\varepsilon_E^s \alpha) \leq \varepsilon_E.$$

A key observation made by Schmidt is that the interval

$$\varepsilon_E^{-1} < x \leq \varepsilon_E$$

may be partitioned into $2t$ intervals $u^{j-1} < x \leq u^j$ with $-t < j \leq t$. Using Schmidt's notation, put $Z_1(\mathfrak{a}, j, X)$ for the number of non-zero $\alpha \in \mathcal{O}_E \cap \mathfrak{a}$ satisfying $|\alpha\overline{\alpha}| \leq X\mathfrak{N}(\mathfrak{a})$ and $u^{j-1} < \psi(\alpha) \leq u^j$. Lemma 6 in [17] gives the estimate

$$Z_1(\mathfrak{a}, j, X) = \frac{2R_E X}{t_E \cdot E^{1/2}} + O\left(\frac{X^{1/2}\mathfrak{N}(\mathfrak{a})^{1/2}}{\lambda_1(\mathfrak{a}, j)}\right). \tag{7.9}$$

Schmidt's Lemma 8 provides a key estimate, namely

$$\sum_{j=1-t}^{t} \lambda_1(\mathfrak{a}, j)^{-1} = \left(\mathfrak{N}(\mathfrak{C}_\mathfrak{a})\mathfrak{N}(\mathfrak{a})\right)^{-1/2}, \tag{7.10}$$

where $\mathfrak{C}_\mathfrak{a}$ is the ideal class containing $\mathfrak{a}$.

We note that for any positive integer k and integer x that

$$\sum_{\substack{c|k \\ c \leq \sqrt{k}}} \left(\frac{x}{c}\right) = \sum_{\substack{c|k \\ c \geq \sqrt{k}}} \left(\frac{x}{c}\right).$$

Using this observation, we may essentially apply Dirichlet's hyperbola trick to (7.1) to obtain

$$\mathcal{S}_h(X) = \mathcal{S}_h^{(1)}(X) + 2 \sideset{}{^\sharp}\sum_{2 \le c \le X^{1/2}} \sideset{}{^*}\sum_{\substack{-Xh_0 < h(x,y) < 0 \\ h(x,y) \equiv 0 \pmod c \\ h_2 x \equiv h_1 y \pmod{h_0}}} \left(\frac{x}{c}\right), \tag{7.11}$$

where the sum $\sum^\sharp$ in (7.11) refers to the restriction to square-free c. Denote this sum by $\mathcal{S}_h^\sharp(X)$.

To proceed, we note that, by Proposition 5.9, we may find a binary quadratic form $\mathfrak{h}$ such that

$$\{h(x,y) : (x,y) \in \mathbb{Z}^2, h_2 x - h_1 y \equiv 0 \pmod{h_0}\} = \{h_0 \cdot \mathfrak{h}(x,y) : (x,y) \in \mathbb{Z}^2\}. \tag{7.12}$$

After replacing h with $\mathfrak{h}$, the congruence condition modulo h_0 then disappears, and the sum $\mathcal{S}_h^\sharp(X)$ becomes

$$\mathcal{S}_h^\sharp(X) = \sideset{}{^\sharp}\sum_{2 \le c \le X^{1/2}} \sideset{}{^*}\sum_{\substack{-X < \mathfrak{h}(x,y) < 0 \\ \mathfrak{h}(x,y) \equiv 0 \pmod c}} \left(\frac{\ell(x,y)}{c}\right) \tag{7.13}$$

for some primitive linear form ℓ.

We break the inner sum of (7.13) into a summation over values b of ℓ, and then summing values of x, y on the line defined by $\ell(x,y) = b$. By the preceding discussion, x, y are constrained in $2t$ domains each with diameter $O\left(X^{1/2}\mathfrak{N}(\mathfrak{C}_\mathfrak{h})^{-1/2}\right)$ (see Lemma 6 and Lemma 11 in [17]) . In particular, we have

$$\begin{aligned}
\sideset{}{^*}\sum_{\substack{-X < \mathfrak{h}(x,y) < 0 \\ \mathfrak{h}(x,y) \equiv 0 \pmod c}} \left(\frac{\ell(x,y)}{c}\right) &= \sum_{b \ll X^{1/2}\mathfrak{N}(\mathfrak{C}_\mathfrak{h})^{-1/2}} \left(\frac{b}{c}\right) \sum_{\substack{(x,y) \in \mathfrak{D}(X) \\ \ell(x,y) = b \\ \mathfrak{h}(x,y) \equiv 0 \pmod c}} 1 \\
&\ll \sqrt{c} \log c \cdot \left(\frac{X^{1/2}}{c\mathfrak{N}(\mathfrak{C}_\mathfrak{h})^{1/2}} + O(1)\right) 2^{\omega(c)} \\
&\ll_\varepsilon X^{1/2}\mathfrak{N}(\mathfrak{C}_\mathfrak{h})^{-1/2} c^{-1/2+\varepsilon}
\end{aligned}$$

Feeding this back into (7.11), and noting that c only runs over norms of $\mathbb{Q}(\sqrt{\Delta(h)})$, and such integers are bounded by $O((X/\Delta(h))^{1/2})$, we see that

$$\begin{aligned}
\mathcal{S}_h^\sharp(X) &\ll_\varepsilon X^{1/2}\mathfrak{N}(\mathfrak{C}_\mathfrak{h})^{-1/2} \sum_{\substack{c \le X^{1/2} \\ c \text{ a norm in } \mathbb{Q}(\sqrt{\Delta(h)})}} c^{-1/2+\varepsilon} \\
&\ll_\varepsilon X^{1/2}\mathfrak{N}(\mathfrak{C}_\mathfrak{h})^{-1/2} \cdot \frac{X^{1/4+\varepsilon}}{\Delta(h)^{1/2}} \\
&= O_\varepsilon\left(X^{3/4+\varepsilon}(\mathfrak{N}(\mathfrak{C}_\mathfrak{h})\Delta(h))^{-1/2}\right).
\end{aligned}$$

The sum corresponding to the value $c = 1$, which we denoted by $\mathcal{S}_h^{(1)}(X)$ in (7.11), is readily seen to be equal to

$$\mathcal{S}_h^{(1)}(X) = \#\{(x,y) \in \mathfrak{D} : -Xh_0 < h(x,y) < 0, h(x,y) \equiv 0 \pmod{h_0}\}$$

where $\mathfrak{D}$ denotes a fundamental domain of the action of the unit group on the ring of integers of $\mathbb{Q}(\sqrt{\Delta(h)})$. By the proof of Theorem 1 in [17], and putting $E = \Delta(h)$, we see that

$$\mathcal{S}_h^{(1)}(X) = \frac{2XR_E}{\sqrt{E}} + O\left((X(\log X)R_E)^{1/2}\right). \tag{7.14}$$

This leads to the following conclusion:

Lemma 7.4. *Let h be an ntegral binary quadratic form with discriminant $E > 0$, and let R_E be the regulator of $\mathbb{Q}(\sqrt{E})$. Then*

$$\mathcal{S}_h(X) = \frac{2XR_E}{\sqrt{E}} + O_\varepsilon\left((X(\log X)R_E)^{1/2} + X^{3/4+\varepsilon}(R_E/E)^{1/2}\right)$$

8. Proof of Theorem 1.1

By Proposition 7.3, it suffices to give $N(X;Y)$ and $\mathcal{N}(X;XY^{-1})$ for any $0<Y<X$. We shall do so for $Y=X^{2/3}\exp\left(\frac{-4\log X}{\log\log X}\right)$.

By Lemma 6.4, we see that the error term in $N_f(X/D)$ can be taken to be

$$O\left(\frac{X^{1/2}}{(\nu_2 D)^{1/2}}\right),$$

where ν_2 is the leading coefficient of $\nu(f)$ is the smallest positive intege representable by $\nu(f)$. Lemma 4.2 then shows that

$$\sum_{D\leq Y}^{\dagger}\frac{X^{1/2}}{(\nu_2 D)^{1/2}}=O\left(X\exp\left(\frac{-2\log X}{\log\log X}\right)\right). \tag{8.1}$$

However, we note that the set of possible classes for $\nu(f)$, by Proposition 5.9, is not all possible classes of discriminant $-D$ but rather restricted to those which are 4-th powers in the class group of forms of discriminant $-D$. For each 4-th power, we then need to multiply by the size of the 4-torsion subgroup of the group of forms of discriminant $-D$ to recover all possible classes of f. The size of the 4-torsion subgroup is at most the square of the size of the 2-torsion subgroup, which by genus theory is at most $2^{\omega(D)}$, where $\omega(\cdot)$ is the number of distinct prime divisor function. Therefore, we need to multiply (8.1) by the largest possible size of the divisor function of a positive integer smaller than X, which is of size $O\left(X^{2/\log\log X}\right)$. We chose our Y to balance this contribution, so that the overall contribution is $O(X)$.

Next, we need to sum over the complementary error terms coming from Lemma 7.4. Examining Schmidt's proof of Lemma 9 in [17] reveals that the geometry of the fundamental domain $\mathfrak{D}$ is affected by the norm of the ideal associated to the norm form h, and we see that the first error term in Lemma 7.4 can be handled as

$$\begin{aligned}\sum_{E\leq Z}\frac{(X(\log X)h_2(E)R_E)^{1/2}}{E^{1/2}}&\leq (X\log X)^{1/2}\left(\sum_{E\leq Z}E^{-1}\right)^{1/2}\left(\sum_{E\leq Z}h_2(E)R_E\right)^{1/2}\quad\text{by Cauchy-Schwarz}\\&\ll (X\log X)^{1/2}(\log Z)^{1/2}Z^{3/4}.\end{aligned}$$

Taking $Z=X^{1/3+1/100}>X^{1/3}\exp\left(4(\log X)(\log\log X)^{-1}\right)$, we see that this gives a negligible contribution.

Now we sum the second error term in Lemma 7.4. Similarly, summing over the classes of discriminant E has the effect of giving the inclusion of the square-root of the class number $h_2(E)$, giving the error term

$$O\left(X^{3/4}(\log X)(h_2(E)R_E/E)^{1/2}\right).$$

Summing, we obtain

$$\begin{aligned}\sum_{E\leq Z}\frac{X^{3/4}\log X}{E^{3/4}}\left(\frac{h_2(E)R_E}{E}\right)^{1/2}&\leq X^{3/4}\log X\left(\sum_{E\leq Z}E^{-1}\right)^{1/2}\left(\sum_{E\leq Z}\frac{h_2(E)R_E}{E^{3/2}}\right)^{1/2}\\&\ll X^{3/4}\log X\cdot\log Z,\end{aligned}$$

by Cauchy-Schwarz and partial summation. Again, by taking $Z=X^{1/3+1/100}$ say, we easily obtain an acceptable error term.

It thus remains to sum over the main terms in Theorem 6.1.

Remark 8.1. Theorem 1.1 is stated with an error term which is $O(X)$, and this is likely not removable. Indeed, akin to the problem of estimating the sum over the divisor function there is likely a secondary main term of size exactly X. Examining the estimation of the summation of errors from Lemmas 6.4 and 7.4 above we see that both cases in fact do give power-saving error terms with a suitable choice of Y, it is in principle possible to obtain the secondary main term. We wish to return to this problem in future work.

8.1. **Summing the main terms.** We now consider the two sums

$$\sum_{\substack{D\leq Y\\ D\equiv 3 \pmod 4\\ w\in W_2^*(-D)}} \frac{\pi X}{3n_w D^{3/2}} \tag{8.2}$$

and

$$\sum_{\substack{D\leq Y\\ D\equiv 0 \pmod 4\\ w\in W_2^*(-D)}} \frac{4\pi X}{3n_w D^{3/2}}, \tag{8.3}$$

where n_w is given as in Proposition 2.6. Put

$$h_2^\sharp(-D) = \sum_{\substack{w\in W_2^*(-D)\\ n_w=1}} 1.$$

We shall then prove the following:

Lemma 8.2. *The equalities*

$$\sum_{\substack{D\leq Y\\ D\equiv 3 \pmod 4}} \frac{h_2^\sharp(-D)\pi X}{3D^{3/2}} = \sum_{\substack{D\leq Y\\ D\equiv 3 \pmod 4}} \frac{h_2(-D)\pi X}{3D^{3/2}} + O(Y\log Y) \tag{8.4}$$

and

$$\sum_{\substack{D\leq Y\\ D\equiv 0 \pmod 4}} \frac{h_2^\sharp(-D)4\pi X}{3D^{3/2}} = \sum_{\substack{D\leq Y\\ D\equiv 0 \pmod 4}} \frac{h_2(-D)4\pi X}{3D^{3/2}} + O(Y\log Y). \tag{8.5}$$

Proof. We have that $n_w > 2$ if and only if $w = [x^2+xy+y^2]_{\mathbb{Z}}$, and it is not possible for a positive definite binary quadratic form f to be opaque. Therefore it suffices to count the number of ambiguous classes with $D \leq Y$ and to show that the number of such classes is small. This requires the estimation of the sum

$$\sum_{D\leq Y} 2^{\omega(D)}.$$

We use the fact that $2^{\omega(n)} = \sum_{d|n}\mu^2(d)$ to obtain

$$\begin{aligned}\sum_{D\leq Y} 2^{\omega(D)} &= Y\sum_{d\leq Y}\mu^2(d)\left(\frac{1}{d}+O(1)\right)\\ &= \frac{8}{27\pi^2}Y\log Y + O(Y).\end{aligned}$$

Hence the number of ambiguous classes with $D \leq Y$ is $O(Y\log Y)$, as desired. □

Lemma 8.2 shows that it suffices to estimate the sum

$$\sum_{\substack{D\leq Y\\ D\equiv 3 \pmod 4}} \frac{h_2(-D)\pi X}{3D^{3/2}} + \sum_{\substack{D\leq Y\\ D\equiv 0 \pmod 4}} \frac{h_2(-D)4\pi X}{3D^{3/2}}. \tag{8.6}$$

We can evaluate (8.6) via Proposition 4.1. Indeed, we note that

$$\begin{aligned}\sum_{\substack{D\leq Y\\ D\equiv 0 \pmod 4}} \frac{h_2(-D)}{D^{3/2}} &= Y^{-3/2}\sum_{\substack{D\leq Y\\ D\equiv 0 \pmod 4}} h_2(-D) + \frac{3}{2}\int_1^Y t^{-5/2}\left(\sum_{\substack{D\leq t\\ D\equiv 0 \pmod 4}} h_2(-D)\right)dt\\ &= \frac{3}{2}\int_1^Y \frac{\pi}{42\zeta(3)}t^{-1}dt + O(1)\\ &= \frac{\pi\log Y}{28\zeta(3)} + O(1).\end{aligned}$$

It thus follows that

$$\sum_{\substack{D\leq Y\\ D\equiv 0 \pmod 4}} \frac{h_2(-D)4\pi X}{3D^{3/2}} = \frac{\pi^2}{21\zeta(3)} X\log Y + O(X). \tag{8.7}$$

Similarly, we evaluate

$$\begin{aligned}\sum_{\substack{D\leq Y\\ D\equiv 3 \pmod 4}} \frac{h_2(-D)}{D^{3/2}} &= Y^{-3/2}\sum_{\substack{D\leq Y\\ D\equiv 3 \pmod 4}} h_2(-D) + \frac{3}{2}\int_1^Y t^{-5/2}\left(\sum_{\substack{D\leq t\\ D\equiv 3 \pmod 4}} h_2(-D)\right)dt\\ &= \frac{3}{2}\int_1^Y \frac{2\pi}{63\zeta(3)} t^{-1}dt + O(1)\\ &= \frac{\pi}{21\zeta(3)}\log Y + O(1),\end{aligned}$$

whence

$$\sum_{\substack{D\leq Y\\ D\equiv 3 \pmod 4}} \frac{h_2(-D)\pi X}{3D^{3/2}} = \frac{\pi^2}{21\zeta(3)} X^{1/3}\log Y + O(X). \tag{8.8}$$

Thus, (8.6) evaluates to

$$\frac{2\pi^2}{21\zeta(3)} X\log Y + O(X). \tag{8.9}$$

Setting $Y = X^{2/3}\exp\left(-4(\log X)(\log\log X)^{-1}\right)$, we obtain the term

$$\frac{4\pi^2}{63\zeta(3)} X\log X + O(X). \tag{8.10}$$

We must perform the same analysis for summing over classes of real quadratic forms. Due to the similarity in calculation, we note that the analogue of (8.6) is

$$\sum_{\substack{E\leq Z\\ E\equiv 0 \pmod 4}} \frac{4Xh_2(E)R_E}{3E^{3/2}} + \sum_{\substack{E\leq Z\\ E\equiv 1 \pmod 4}} \frac{Xh_2(E)R_E}{3E^{3/2}}, \tag{8.11}$$

and by the second half of Proposition 4.1, we obtain the same conclusion as in the positive definite case. It thus follows that

$$N(X) = N(X;Y) + \mathcal{N}(X;XY^{-1}) = \left(\frac{4\pi^2}{63\zeta(3)} + \frac{2\pi^2}{63\zeta(3)}\right) X\log X + O(X) = \frac{2\pi^2}{21\zeta(3)} X\log X + O(X). \tag{8.12}$$

Finally, by replacing X with $3X^{1/3}/\sqrt[3]{4}$, we complete the proof of Theorem 1.1.

9. Proof of Theorem 1.2

Compared with the proof of Theorem 1.1, the proof of Theorem 1.2 is much simpler, since the reduced classes of reducible binary quadratic forms are particularly simple. Here we find that a typical reducible and reduced binary quadratic form takes the shape

$$f(x,y) = \alpha x^2 + \beta xy, \gcd(\alpha,\beta) = 1. \tag{9.1}$$

This then implies that the lattice $\mathcal{L}_{f,\alpha}$ takes a particularly simple shape, namely

$$\mathcal{L}_{f,\alpha} = \{(x,y)\in\mathbb{Z}^2 : y \equiv 0 \pmod{4\alpha^3}\}.$$

We thus replace B with $4\alpha^3 B$, so that the generic element $F\in\mathcal{V}_f(\mathbb{Z})$ takes the form

$$F(x,y) = Ax^4 + 4\alpha^3 Bx^3 y + 6\alpha^2\beta Bx^2y^2 + 4\alpha\beta^2 Bxy^3 + \beta^3 By^4. \tag{9.2}$$

It then follows that $\mathcal{I}(F)$ is given by

$$\mathcal{I}(F) = 4B(4\alpha^7 B - \beta A).$$

We then have:

Lemma 9.1. *We have*

$$N_f(X) = \frac{X^{1/3}}{3\beta^3}\log\left(\frac{X^{1/3}}{3\beta^2}\right) + (2\gamma - 1)\frac{X^{1/3}}{3\beta^3} + O\left(\frac{X^{1/6}}{\beta}\right).$$

Proof. By symmetry, either B or $(4\alpha^7 B - \beta A)$ is less than $X^{1/3}/(12\beta^2)$ in absolute value. We shall assume that $B, 4\alpha^7 B - \beta \geq 1$. For convenience, we shall put $Y = X^{1/3}/(12\beta^2)$. We then look at the three sums

$$S_1(X) = \sum_{m \leq Y^{1/2}} \sum_{\substack{n \leq Y/m \\ n \equiv 4\alpha^7 m \pmod{\beta}}} 1,$$

$$S_2(X) = \sum_{m \leq Y^{1/2}} \sum_{\substack{n \leq Y/m \\ m \equiv 4\alpha^7 n \pmod{\beta}}} 1$$

and

$$S_3(X) = \sum_{m \leq Y^{1/2}} \sum_{\substack{n \leq Y^{1/2} \\ n \equiv 4\alpha^7 m \pmod{\beta}}} 1.$$

It is then clear that

$$N_f(X) = S_1(X) + S_2(X) - S_3(X).$$

We evaluate $S_1(X)$. The inner sum is equal to $\frac{Y}{\beta m} + O(1)$. Thus, we have

$$S_1(X) = \frac{1}{2\beta}\left(Y\log Y + 2\gamma Y\right) + O(Y^{1/2}).$$

Here γ is the Euler-Mascheroni constant. The evaluation of $S_2(X)$ is the same, and we have that $S_1(X) = S_2(X) + O(Y^{1/2})$. It is easy to see that

$$S_3(X) = \frac{Y}{\beta} + O(Y^{1/2}).$$

It thus follows that

$$\begin{aligned} N_f(X) &= \frac{Y\log Y + (2\gamma - 1)Y}{\beta} + O(Y^{1/2}) \\ &= \frac{X^{1/3}\log(X^{1/3}/(12\beta^2) + (2\gamma - 1)X^{1/3}}{12\beta^3} + O\left(\frac{X^{1/6}}{\beta}\right). \end{aligned}$$

Multiplying by 4 to account for the signs of B and $4\alpha^7 - \beta A$, we obtain the result. □

We may now prove Theorem 1.2.

Proof of Theorem 1.2. The error term in the estimate for $N_f(X)$ provided by Lemma 9.1 is sufficiently sharp that we may evaluate the sum directly using the class number formula given by Proposition 4.3. We are then left to evaluate the sum

$$\frac{1}{3}\sum_{n \leq X^{1/6}} \phi(n)\left(\frac{X^{1/3}}{n^3}\log X - \frac{X^{1/3}}{n^3}\log(12n^2) + (2\gamma - 1)\frac{X^{1/3}}{n^3} + O\left(\frac{X^{1/6}}{n}\right)\right).$$

We have the well-known identity

$$\sum_{n \geq 1}\frac{\phi(n)}{n^s} = \frac{\zeta(s-1)}{\zeta(s)}, \Re(s) > 2,$$

Hence

$$\frac{X^{1/3}}{3}\left((\log X)/3 + 2\gamma - 1\right)\sum_{n \leq X^{1/6}}\frac{\phi(n)}{n^3} = \frac{X^{1/3}\left((\log X)/3 + 2\gamma - 1\right)}{3}\left(\frac{\zeta(2)}{\zeta(3)} + O(X^{-1/6})\right).$$

By partial summation, we see that

$$\frac{X^{1/3}}{3}\sum_{n \leq X^{1/6}}\frac{\phi(n)\log(12n^2)}{n^3} = O(X^{1/3}).$$

Since $\phi(n) \leq n-1$ for all positive integers n, it follows that

$$\sum_{n \leq X^{1/6}} \frac{\phi(n)}{n} = O(X^{1/6}).$$

Finally, we need to address the issue of ambiguous and opaque classes. By Proposition 4.12 in [20], we see that the number of classes of discriminant n^2 which are opaque or ambiguous is at most $2^{\omega(n)+1}$. We then note that

$$\sum_{n \leq X^{1/6}} \frac{2^{\omega(n)}}{n^3} = O(1).$$

Therefore, very few classes are ambiguous or opaque, and the proof is complete. □

10. Proof of Theorems 1.3 - 1.5

10.1. **Proof of Theorem 1.3.** Let F be given by (1.1). Consider its *cubic resolvent polynomial* given by

$$R_F(x) = a_4^3X^3 - a_4^2a_2X^2 + a_4(a_3a_1 - 4a_4a_0)X - (a_3^2a_0 + a_4a_1^2 - 4a_4a_2a_0).$$

It is well-known that for irreducible F, $R_F(x)$ has a rational root if and only if $\mathrm{Gal}(F)$ is isomorphic to a subgroup of D_4. In [24] we showed that $R_F(x)$ has a rational root if and only if F has a rational *Cremona covariant*. For $F \in \mathcal{V}_f(\mathbb{Z})$, f is a rational Cremona covariant of F, hence $R_F(x)$ has a rational root and $\mathrm{Gal}(F)$ is D_4, C_4, or V_4.

We first suppose that $-\Delta(f)$ is not a square. Suppose that $F \in \mathcal{V}_f(\mathbb{Z})$. If $\Delta(F)$ is itself a square then $\mathrm{Gal}(F)$ cannot be isomorphic to C_4, so we assume that $\Delta(F) \neq \square$. $R_F(x)$ has a unique root $r_F \in \mathbb{Q}$ precisely when $\Delta(F) \neq \square$, in which case we define

$$\theta_1(F) = (a_3^2 - 4a_4(a_2 - r_Fa_4))\Delta(F) \text{ and } \theta_2(F) = a_4(r_F^2a_4 - 4a_0)\Delta(F).$$

It is well-known (see [9]) that $\mathrm{Gal}(F) \cong C_4$ precisely when $\Delta(F) \neq \square$ and $\theta_1(F), \theta_2(F)$ are rational squares. Writing $a_4, \cdots, a_0$ as in (2.6) we find that

$$\theta_1(F) = \frac{-\Delta(f)^3(\alpha B^2 - 4\beta AB + 16\gamma A^2)^4}{16\alpha^{10}}$$

and

$$\theta_2(F) = \frac{-\Delta(f)^3(\alpha B^2 - 4\beta AB + 16\gamma A^2)^4}{64\alpha^{12}}.$$

Thus, it is apparent that both $\theta_1(F), \theta_2(F)$ are squares modulo $-\Delta(f)$. Since $-\Delta(f)$ is not a square by assumption, neither are $\theta_1(F), \theta_2(F)$.

Now suppose that $-\Delta(f)$ is a square. By the same argument as above, we see that whenever F is irreducible and $\Delta(F)$ is not a square, we have that $\mathrm{Gal}(F) \cong C_4$. It thus remains to show that whenever $\Delta(F)$ is a square, that F is reducible.

10.2. **Proof of Theorem 1.4.** Since all elements in $\mathcal{V}_4^{(0)}(\mathbb{R})$ lie in a single $\mathrm{GL}_2(\mathbb{R})$-orbit, it suffices to consider the statement for a single element in $\mathcal{V}_4^{(0)}(\mathbb{R})$. We take

$$F = xy(x^2 - y^2).$$

It is easily verified that $J(F) = 0$ and F is totally real. Moreover, we have $I(F) = 3$. We compute

$$F_6(x, y) = x^6 - 5x^4y^2 - 5x^2y^4 + y^6 = (x^2 + y^2)(x^2 - 2xy - y^2)(x^2 + 2xy - y^2).$$

We then find that

$$G_F = (x^2 - 2xy - y^2)(x^2 + 2xy - y^2).$$

Note that $J(G_F) = 0$ and G_F is also totally real. We then find that

$$(G_F)_6(x, y) = 2(16 - 144)x^5 - 2(16 - 144)xy^5 = 256xy(x^2 + y^2)(x^2 - y^2).$$

Therefore,

$$G_{G_F} = 256xy(x^2 - y^2),$$

which is proportional to F as claimed.

10.3. **Proof of Theorem 1.5.** Observe that if $-\Delta(f)$ is a square, then it is necessarily even. Hence for each such f we have, by Theorem 6.1,

$$N_f(X) = \frac{4\pi}{3(4n^2)^{3/2}} X^{1/3} + O\left(\frac{X^{1/6}}{n}\right) = \frac{\pi}{6n^3} X^{1/3} + O\left(\frac{X^{1/6}}{n}\right).$$

The error term from Theorem 6.1 is not sufficient for our purposes. However, by the same argument given in Section 8 we may first consider all classes of $\nu(f)$, then multiply by the size of the 4-torsion subgroup of the class group. This wins us an extra factor of ν_2, the x^2-coefficient of $\nu(f)$, in the denominator.

We now consider reduced classes of $\nu(f)$, that is, the set

$$R'(X) = \{(a,b,c) \in \mathbb{Z}^3 : |b| \le a \le c, ac > 0, b^2 - ac = -n^2, 1 \le n < X^{1/6}/2\}.$$

Let $\varepsilon > 0$. We consider the subset $R''(X)$ of $R'(X)$ with $a \le n^\varepsilon$. For each $n \in [1, X^{1/6}/2)$ put $h_2^\sharp(-4n^2)$ for the classes counted by $h_2(-4n^2)$ which correspond to an element in $R''(X)$. There are n^ε choices for a, $2a = O(n^\varepsilon)$ choices for b, and $d(n^2+b^2) = O_\varepsilon(n^\varepsilon)$ choices for c. Finally, for each such n there are at most $O_\varepsilon(n^\varepsilon)$ elements in the 4-torsion subgroup of the Picard group of $\mathcal{O}_{-4n^2}$. Thus, adjusting ε if necessary, we have $h_2^\star(-4n^2) = O_\varepsilon(n^\varepsilon)$. Otherwise we have the bound $h_2(-4n^2) \ll n \log\log n$. Combining these bounds we obtain that summing the error term over all classes counted $R'(X)$ gives an error term of

$$\sum_{n \le X^{1/6}/2} O_\varepsilon\left(\frac{X^{1/6}}{n^{1-\varepsilon}} + \frac{X^{1/6}\log\log n}{n^\varepsilon}\right) = O_\varepsilon\left(X^{1/3-\varepsilon}\right).$$

We now sum over the main term. The sum over all such $\mathrm{SL}_2(\mathbb{Z})$-classes of f gives the sum

$$\sum_{n \le X^{1/6}/2} \frac{h_2(-4n^2)\pi}{6n^3} X^{1/3} = X^{1/3} \frac{\pi}{6} \frac{7\zeta(2)}{8\beta(3)} = X^{1/3} \frac{\pi \cdot \pi^2 \cdot 28}{6 \cdot 6 \cdot \pi^3} = \frac{7X^{1/3}}{9}.$$

It thus follows that

$$N_{C_4}(X) = \frac{7X^{1/3}}{9} + O_\varepsilon\left(X^{1/3-\varepsilon}\right),$$

and the estimate for $M_{C_4}(X)$ follows by replacing X with $4X/27$.

Finally, we show that all irreducible elements in $\mathcal{V}_f(\mathbb{Z})$ have Galois group C_4. We have already ruled out D_4, so it suffices to show that all elements $F \in \mathcal{V}_f(\mathbb{Z})$ with square discriminant are in fact reducible. Since $-\Delta(f) = \square$, it follows that f is $\mathrm{GL}_2(\mathbb{Q})$-equivalent to x^2+y^2. Therefore, F is $\mathrm{GL}_2(\mathbb{Q})$-equivalent to a form of the shape

$$\mathcal{F} = Ax^4 + Bx^3y - 6Ax^2y^2 - Bxy^3 + Ay^4, A, B \in \mathbb{Q}.$$

By the same argument as in the proof of Theorem 1.4, we find that

$$G_{\mathcal{F}}(x,y) = (16A^2+B^2)(Bx^4 - 16Ax^3y - 6Bx^2y^2 + 16Axy^3 + By^4).$$

We now suppose that $\Delta(F)$ is a square, which is a necessary condition for $\mathrm{Gal}(F) \cong V_4$. By the proof of Theorem 1.1 in [20], we find that $G_{\mathcal{F}}$ is necessarily reducible. Moreover $\Delta(G_{\mathcal{F}})$ is a square, so then it follows that $G_{G_{\mathcal{F}}}$ is necessarily reducible. But $G_{G_{\mathcal{F}}}$ is proportional over $\mathbb{Q}$ to F. Hence F is reducible, as claimed.

Appendix: A Coordinate-free Perspective on the Hensel Lifts
By Erick Knight

This appendix is an alternative perspective on the discussion in Section 5; in particular it is a coordinate-free perspective on Proposition 5.9. In this discussion, we will restrict ourselves to considering the class group of a single field K.

To fix some notation, let K be a quadratic extension of $\mathbb{Q}$, with ring of integers $\mathcal{O}_K$. Additionally, let p be a prime of $\mathbb{Z}$ that splits in K, and write $(p) = \mathfrak{p}_1\mathfrak{p}_2$. We will denote by $\mathcal{O}_{K,\mathfrak{p}_i}$ to be the $\mathfrak{p}_i$-adic completion of $\mathcal{O}_K$. Because p splits in K, the natural inclusion $\mathbb{Z}_p \hookrightarrow \mathcal{O}_{K,\mathfrak{p}_i}$ is an isomorphism. Additionally, one has that $\mathcal{O}_K \otimes_{\mathbb{Z}} \mathbb{Z}_p \cong \mathcal{O}_{K,\mathfrak{p}_1} \oplus \mathcal{O}_{K,\mathfrak{p}_2}$. Using these isomorphisms, one has that the norm form $N_{K/\mathbb{Q}} : \mathcal{O}_K \otimes_{\mathbb{Z}} \mathbb{Z}_p \to \mathbb{Z}_p$ is just the composition $\mathcal{O}_K \otimes_{\mathbb{Z}} \mathbb{Z}_p \to \mathcal{O}_{K,\mathfrak{p}_1} \oplus \mathcal{O}_{K,\mathfrak{p}_2} \to \mathbb{Z}_p \oplus \mathbb{Z}_p \to \mathbb{Z}_p$ where the first two maps are just the isomorphisms mentioned earlier, and the last map is just $(x,y) \to xy$.

Now, let f be a quadratic form such that $[f]$ is equal to $[\mathfrak{p}_1]$. This means that there is an identification of $\mathbb{Z}^2$ with $\mathfrak{p}_1$ such that f is equal to the function $\frac{N_{K/\mathbb{Q}}(\cdot)}{p}$ in this basis. Tensoring up to $\mathbb{Z}_p$, we get that $\mathfrak{p}_1 \otimes_{\mathbb{Z}} \mathbb{Z}_p \cong p\mathcal{O}_{K,\mathfrak{p}_1} \oplus \mathcal{O}_{K,\mathfrak{p}_2}$ as an ideal in $\mathcal{O}_K \otimes_{\mathbb{Z}} \mathbb{Z}_p$, and f sends an element (α, β) of $p\mathcal{O}_{K,\mathfrak{p}_1} \oplus \mathcal{O}_{K,\mathfrak{p}_2}$ to $\alpha\beta/p$.

This means that the lattices constructed in Section 5 are given by taking the intersection $\mathcal{O}_K \cap p^{s+1}\mathcal{O}_{K,\mathfrak{p}_1} \oplus \mathcal{O}_{K,\mathfrak{p}_2}$ inside of $\mathcal{O}_K \otimes_{\mathbb{Z}} \mathbb{Z}_p$. But this is just $\mathfrak{p}_1^{s+1}$, as can be seen from the Chinese remainder theorem. Moreover, the form $g_{2,k}$ is just given by restriciting f to this lattice and then dividing by p^s, which means that $[g_{2,k}]$ is equal to the class of $[\mathfrak{p}_1^{s+1}]$, which is what was wanted.

References

[1] M. A. Bennett, S. R. Dahmen, *Klein forms and the generalized superelliptic equation*, Ann. of. Math **177** (2013), 171-239.
[2] M. Bhargava, *Higher composition laws I: A new view on Gauss composition, and quadratic generalizations*, Ann. of. Math, **159** (2004), 217-250.
[3] M. Bhargava, A. Shankar, *Binary quartic forms having bounded invariants, and the boundedness of the average rank of elliptic curves*, Ann. of Math **181** (2015), 191-242.
[4] M. Bhargava, A. Shankar, J. Tsimerman, *On the Davenport-Heilbronn theorems and second order terms*, Invent. math., **193** (2013), 193-439
[5] M . Bhargava, A. Shnidman, *On the number of cubic orders of bounded discriminant having automorphism group C_3, and related problems*, Algebra and Number Theory, (1) **8** (2014), 53-88.
[6] V. Blomer, A. Granville, *Estimates for representation numbers of binary quadratic forms*, Duke. Math. J (2) **135** (2006), 261-302.
[7] A. Borel, Harish-Chandra, *Arithmetic subgroups of algebraic groups*, Ann. of. Math **75** (1962), 485-535.
[8] D. Buell, *Binary Quadratic Forms - Classical Theory and Modern Computations*, Springer-Verlag New York, 1989.
[9] K. Conrad, *Galois groups of cubics and quartics (not in characteristic 2)* `http://www.math.uconn.edu/~kconrad/blurbs/galoistheory/cubicquartic.pdf`, retrieved 30 Oct 2015.
[10] J. Cremona, *Reduction of binary cubic and quartic forms*, LMS J. Comput. Math. **2** (1999), 64-94.
[11] H. Davenport, *On a principle of Lipschitz*, J. London Math.Soc., **26** (1951), 179-183.
[12] H. Davenport, *On the class-number of binary cubic forms*, I, J. London Math.Soc. **26** (1951), 183-192.
[13] C. F. Gauss, *Disquisitiones Arithmeticae*, 1801.
[14] D. R. Heath-Brown, *A new form of the circle method, and its application to quadratic forms*, J. Reine Angew. Math **481** (1996), 149-206.
[15] F. Mertens, *Ueber einige asymptotische Gesetze der Zahlentheorie*, J. reine angew Math. **77** (1874), 289-338.
[16] J. Morales, The classification of pairs of binary quadratic forms. Acta Arith (2) **59** (1991),105-121.
[17] W. M. Schmidt, *Northcott's theorem on heights II. The quadratic case*, Acta. Arith. **70** (1995), 343-375.
[18] C. L. Siegel, *The average measure of quadratic forms with given determinant and signature*, Ann. of. Math (4) **45** (1944), 667-685.
[19] C. L. Stewart, *On the number of solutions of polynomial congruences and Thue equations*, J. Amer. Math. Soc (4) **4** (1991), 793-835.
[20] C. Tsang, S. Y. Xiao, *Binary quartic forms with bounded invariants and small Galois groups*, to appear in Pac. J. Math. arXiv:1702.07407 [math.NT].
[21] C. Tsang, S. Y. Xiao, *The number of quartic D_4-fields with monogenic cubic resolvent ordered by conductor*, arXiv:1712.08552 [math.NT].
[22] M. Wood, *Moduli spaces for rings and ideals*, ProQuest LLC, Ann Arbor, MI, 2009, Ph.D. thesis, Princeton University.
[23] M. Wood, *Quartic rings associated to binary quartic forms*, Int. Math. Res. Notices, Volume 2012 Issue 6, 1300-1320.
[24] S. Y. Xiao, *On binary cubic and quartic forms*, to appear in J. Theor. Nombres Bordeaux.

Department of Mathematics, University of Toronto, Bahen Centre, 40 St. George Street, Room 6290, Toronto, Ontario, Canada, M5S 2E4

E-mail address: `syxiao@math.toronto.edu`