Annals of Mathematics, **162** (2005), 1031–1063

The density of discriminants of quartic rings and fields

By Manjul Bhargava

1. Introduction

The primary purpose of this article is to prove the following theorem.

Theorem 1. *Let $N_4^{(i)}(\xi,\eta)$ denote the number of S_4-quartic fields K having $4-2i$ real embeddings such that $\xi < \mathrm{Disc}(K) < \eta$. Then*

$$\begin{array}{llcl}
\text{(a)} & \displaystyle\lim_{X\to\infty}\frac{N_4^{(0)}(0,X)}{X} & = & \displaystyle\frac{1}{48}\prod_p(1+p^{-2}-p^{-3}-p^{-4});\\
\text{(b)} & \displaystyle\lim_{X\to\infty}\frac{N_4^{(1)}(-X,0)}{X} & = & \displaystyle\frac{1}{8}\prod_p(1+p^{-2}-p^{-3}-p^{-4});\\
\text{(c)} & \displaystyle\lim_{X\to\infty}\frac{N_4^{(2)}(0,X)}{X} & = & \displaystyle\frac{1}{16}\prod_p(1+p^{-2}-p^{-3}-p^{-4}).
\end{array}$$

Several further results are obtained as by-products. First, our methods enable us to count all orders in S_4-quartic fields.

Theorem 2. *Let $M_4^{(i)}(\xi,\eta)$ denote the number of quartic orders $\mathcal{O}$ contained in S_4-quartic fields having $4-2i$ real embeddings such that $\xi<\mathrm{Disc}(\mathcal{O})<\eta$. Then*

$$\begin{array}{llcl}
\text{(a)} & \displaystyle\lim_{X\to\infty}\frac{M_4^{(0)}(0,X)}{X} & = & \displaystyle\frac{\zeta(2)^2\zeta(3)}{48\,\zeta(5)};\\
\text{(b)} & \displaystyle\lim_{X\to\infty}\frac{M_4^{(1)}(-X,0)}{X} & = & \displaystyle\frac{\zeta(2)^2\zeta(3)}{8\,\zeta(5)};\\
\text{(c)} & \displaystyle\lim_{X\to\infty}\frac{M_4^{(2)}(0,X)}{X} & = & \displaystyle\frac{\zeta(2)^2\zeta(3)}{16\,\zeta(5)}.
\end{array}$$

Second, the proof of Theorem 1 involves a determination of the densities of various splitting types of primes in S_4-quartic fields. If K is an S_4-quartic field unramified at a prime p, and K_{24} denotes the Galois closure of K, then the

Artin symbol (K_{24}/p) is defined as a conjugacy class in S_4, its values being $\langle e\rangle$, $\langle(12)\rangle$, $\langle(123)\rangle$, $\langle(1234)\rangle$, or $\langle(12)(34)\rangle$, where $\langle x\rangle$ denotes the conjugacy class of x in S_4. It follows from the Chebotarev density theorem that for fixed K and varying p (unramified in K), the values $\langle e\rangle$, $\langle(12)\rangle$, $\langle(123)\rangle$, $\langle(1234)\rangle$, and $\langle(12)(34)\rangle$ occur with relative frequency $1:6:8:6:3$. We prove the following complement to Chebotarev density:

THEOREM 3. *Let p be a fixed prime, and let K run through all S_4-quartic fields in which p does not ramify, the fields being ordered by the size of the discriminants. Then the Artin symbol (K_{24}/p) takes the values $\langle e\rangle$, $\langle(12)\rangle$, $\langle(123)\rangle$, $\langle(1234)\rangle$, and $\langle(12)(34)\rangle$ with relative frequency* $1:6:8:6:3$.

Actually, we do a little more: we determine for each prime p the density of quartic fields K in which p has the various possible ramification types. For instance, it follows from our methods that a proportion of precisely $\frac{(p+1)^2}{p^3+p^2+2p+1}$ of S_4-quartic fields are ramified at p.

Third, Theorem 1 implies that relatively many—in fact, a positive proportion of!—quartic fields do not have full Galois group S_4. Indeed, it was shown by Baily [1], using methods of class field theory, that the number of D_4-quartic fields having absolute discriminant less than X is between c_1X and c_2X for some constants c_1 and c_2. This result was recently refined to an exact asymptotic by Cohen, Diaz y Diaz, and Olivier [7], who showed that the number of such D_4-quartic fields is $\sim cX$, where $c \approx .052326\ldots$. Moreover, it has been shown by Baily [1] and Wong [26] that the contributions from the Galois groups C_4, K_4, and A_4 are negligible in comparison; i.e., the number of quartic extensions having one of these Galois groups and absolute discriminant at most X is $o(X)$ (in fact, $O(X^{\frac{7}{8}+\epsilon})$). In conjunction with these results, Theorem 1 implies:

THEOREM 4. *When ordered by absolute discriminant, a* positive proportion (*approximately* 17.111%) *of quartic fields have associated Galois group D_4. The remaining* 82.889% *of quartic fields have Galois group S_4.*

As noted in [6], this is in stark contrast to the situation for polynomials, since Hilbert showed that 100% of degree n polynomials (in an appropriate sense) have Galois group S_n. Theorem 4 may be broken down by signature. Among the quartic fields having 0, 2, or 4 complex embeddings respectively, the proportions having associated Galois group S_4 are given by: 83.723%, 93.914%, and 66.948% respectively.

Finally, using a duality between quartic fields and 2-class groups of cubic fields, we are able to determine the mean value of the size of the 2-class group of both real and complex cubic fields. More precisely, we prove

THEOREM 5. *For a cubic field F, let $h_2^*(F)$ denote the size of the exponent-2 part of the class group of F. Then*

$$\text{(a)} \quad \lim_{X\to\infty} \frac{\sum_F h_2^*(F)}{\sum_F 1} \;=\; 5/4\,; \tag{1}$$

$$\text{(b)} \quad \lim_{X\to\infty} \frac{\sum_F h_2^*(F)}{\sum_F 1} \;=\; 3/2\,, \tag{2}$$

where the sums range over cubic fields F having discriminants in the ranges $(0,X)$ and $(-X,0)$ respectively.

The theorem implies, in particular, that at least 75% of totally real cubic fields, and at least 50% of complex cubic fields, have odd class number.

It is natural to compare the values 5/4 and 3/2 obtained in our theorem with the corresponding values predicted by the Cohen-Martinet heuristics (the analogues of the Cohen-Lenstra heuristics for noncyclic, higher degree fields). There has been much recent skepticism surrounding these heuristics (even by Cohen-Martinet themselves; see [9]), since at the prime $p = 2$ they do not seem to agree with existing computational data.* In light of this situation, it is interesting to note that our Theorem 5 *agrees exactly* with the (original) prediction of the Cohen-Martinet heuristics [8]. In particular, Theorem 5 is a strong indication that, in the language of [8], the prime $p = 2$ is indeed "good", and the fact that Theorem 5 does not agree well with current computations is due only to the extremely slow convergence of the limits (1) and (2).

The cubic analogues of Theorems 1, 3, and 5 for cubic fields were obtained in the well-known work of Davenport-Heilbronn [15]. Their methods relied heavily on the remarkable discriminant-preserving correspondence between cubic orders and equivalence classes of integral binary cubic forms, established by Delone-Faddeev [16]. It seems, however, that Davenport-Heilbronn were not aware of the work in [16], and derived the same correspondence for maximal orders independently; had they known the general form of the Delone-Faddeev parametrization, it would have been possible for them (using again the results of Davenport [13]) simply to read off the cubic analogue of Theorem 2.† Mean-

*A computation of all real cubic fields of discriminant less than 500000 ([17]) shows that $(\sum_{0<\mathrm{Disc}(F)<500000} h_2^*(F))/(\sum_{0<\mathrm{Disc}(F)<500000} 1)$ equals about 1.09, a good deal less than 5/4; the analogous computation for complex cubic fields of absolute discriminant less than 1000000 ([18]) yields approximately 1.30, a good deal less than 3/2!

†We note the result here, since it seems not to have been stated previously in the literature. Let $M_3(\xi,\eta)$ denote the number of cubic orders $\mathcal{O}$ such that $\xi < \mathrm{Disc}(\mathcal{O}) < \eta$. Then

$$\lim_{X\to\infty} \frac{M_3(0,X)}{X} = \pi^2/72,$$
$$\lim_{X\to\infty} \frac{M_3(-X,0)}{X} = \pi^2/24.$$

while, the cubic analogue of Theorem 4 may be obtained by combining the work of Davenport-Heilbronn [15] with that of Cohn [10].[‡]

An important ingredient that allows us to extend the above cubic results to the quartic case is a parametrization of quartic orders by means of two integral ternary quadratic forms up to the action of $\mathrm{GL}_2(\mathbb{Z}) \times \mathrm{SL}_3(\mathbb{Z})$, which we established in [3]. The proofs of Theorems 1–5 thus reduce to counting integer points in certain 12-dimensional fundamental regions. We carry out this counting in a hands-on manner similar to that of Davenport [13], although another crucial ingredient in our work is a new averaging method which allows us to deal more efficiently with points in the cusps of these fundamental regions. The necessary point-counting is accomplished in Section 2. This counting result, together with the results of [3], immediately yields the asymptotic density of discriminants of pairs (Q, R), where Q is an order in an S_4-quartic field and R is a cubic resolvent of Q. Obtaining Theorems 1–5 from this general density result then requires a sieving process which we carry out in Section 3.

The space of pairs of ternary quadratic forms that we use in this article, as well as the space of binary cubic forms that was used in the work of Davenport-Heilbronn, are both examples of what are known as prehomogeneous vector spaces. A *prehomogeneous vector space* is a pair (G, V), where G is a reductive group and V is a linear representation of G such that $G_{\mathbb{C}}$ has a Zariski open orbit on $V_{\mathbb{C}}$. The concept was introduced by Sato in the 1960's, and a classification of all prehomogeneous vector spaces was given in the work of Sato-Kimura [22], while Sato-Shintani [23] developed a theory of zeta functions associated to these spaces.

The connection between prehomogeneous vector spaces and field extensions was first studied systematically in the beautiful 1992 paper of Wright-Yukie [27]. In that paper, they laid out a program to determine the density of discriminants of number fields of degree up to five by considering adelic versions of Sato-Shintani's zeta functions as developed by Datskovsky and Wright [11] in their work on cubic extensions. Despite looking very promising, the program has not succeeded to date beyond the cubic case, although the global theory of the adelic zeta function in the quartic case was developed in the impressive 1993 treatise of Yukie [28], which led to a conjectural determination of the Euler products appearing in Theorem 1 (see [29]).

The reason that the zeta function method has required such a large amount of work, and has thus presented some related difficulties, is that intrinsic to the zeta function approach is a certain overcounting of quartic extensions. Specifically, even when one wishes to count only quartic field extensions of $\mathbb{Q}$ having, say, Galois group S_4, inherent in the zeta function is a sum over all

[‡]Their work implies that, when ordered by absolute discriminant, 100% of cubic fields have associated Galois group S_3.

"étale extensions" of $\mathbb{Q}$, including the "reducible" extensions that correspond to direct sums of quadratic extensions. These reducible quartic extensions far outnumber the irreducible ones; indeed, the number of reducible quartic extensions of absolute discriminant at most X is asymptotic to $X \log X$, while we show that the number of quartic field extensions of absolute discriminant at most X is only $O(X)$. This overcount results in the Shintani zeta function having a double pole at $s = 1$ rather than a single pole. Removing this double pole, in order to obtain the desired main term, has been the primary difficulty with the zeta function method.

One way our viewpoint differs from the adelic zeta function approach is that we consider *integer orbits* as opposed to rational orbits. This turns out to have a number of significant advantages. First, the use of integer orbits enables us to apply a convenient reduction theory in terms of Siegel sets. Within these Siegel sets, we then determine which regions contain many irreducible points and which do not. We prove that the cusps of the Siegel sets contain most of the reducible points, while the main bodies of the Siegel sets contain most of the irreducible points. These geometric results allow us to separate the irreducible orbits from the reducible ones from the very beginning, so that we may proceed directly to the "irreducible" integer orbits, where geometry-of-numbers methods are applicable. The aforementioned difficulties arising from overcounting are thus bypassed.

A second important advantage of using integer orbits in conjunction with geometry-of-numbers arguments is that the resulting methods are very elementary and the treatment is relatively short. Finally, the use of integer orbits enables us to count not only S_4-quartic fields but also all orders in S_4-quartic fields.

Nevertheless, the adelic zeta function method, if completed in the future, could lead to some interesting results to supplement Theorems 1–5. For example, it may yield functional equations for the zeta function as well as a precise determination of its poles, thus possibly leading to lower bounds on first order error terms in Theorem 1–5. It is also likely that the zeta function methods together with the methods introduced here would lead to even further applications in these and other directions.

We fully expect that the geometric methods introduced in this paper will also prove useful in other contexts. For example, with only slight modifications, the methods of this paper can also be used to derive the density of discriminants of *quintic* orders and fields. These and related results will appear in [4], [5].

We note that, in this paper, we always count quartic (and cubic) number fields up to isomorphism. Another natural way to count number fields is as subfields of a fixed algebraic closure $\bar{\mathbb{Q}}$ of $\mathbb{Q}$. It is easy to see that any isomorphism class of S_4-quartic field corresponds to four conjugate subfields of $\bar{\mathbb{Q}}$, while an isomorphism class of D_4-quartic field corresponds to two conjugate subfields of $\bar{\mathbb{Q}}$. Adopting the latter counting convention would therefore multiply all constants in Theorems 1 and 2 by a factor of four. Moreover, the proportion of S_4-quartic fields in Theorem 4 would then increase to 90.644% (by signature: 91.141%, 96.862%, and 80.202%). Theorems 3 and 5, of course, would remain unchanged.

2. On the class numbers of pairs of ternary quadratic forms

Let $V_{\mathbb{R}}$ denote the space of pairs (A, B) of ternary quadratic forms over the real numbers. We write an element $(A, B) \in V_{\mathbb{R}}$ as a pair of 3×3 symmetric real matrices as follows:

$$2 \cdot (A, B) = \left(\begin{bmatrix} 2a_{11} & a_{12} & a_{13} \\ a_{12} & 2a_{22} & a_{23} \\ a_{13} & a_{23} & 2a_{33} \end{bmatrix}, \begin{bmatrix} 2b_{11} & b_{12} & b_{13} \\ b_{12} & 2b_{22} & b_{23} \\ b_{13} & b_{23} & 2b_{33} \end{bmatrix} \right). \tag{3}$$

Such a pair (A, B) is said to be *integral* if A and B are "integral" quadratic forms, i.e., if $a_{ij}, b_{ij} \in \mathbb{Z}$.

The group $G_{\mathbb{Z}} = \mathrm{GL}_2(\mathbb{Z}) \times \mathrm{SL}_3(\mathbb{Z})$ acts naturally on the space $V_{\mathbb{R}}$. Namely, an element $g_2 \in \mathrm{GL}_2(\mathbb{Z})$ acts by changing the basis of the lattice of forms spanned by (A, B); i.e., if $g_2 = \left(\begin{smallmatrix} r & s \\ t & u \end{smallmatrix}\right)$, then $g_2 \cdot (A, B) = (rA + sB, tA + uB)$. Similarly, an element $g_3 \in \mathrm{SL}_3(\mathbb{Z})$ changes the basis of the three-dimensional space on which the forms A and B take values; i.e., $g_3 \cdot (A, B) = (g_3 A g_3^t, g_3 B g_3^t)$. It is clear that the actions of g_2 and g_3 commute, and that this action of $G_{\mathbb{Z}}$ preserves the lattice $V_{\mathbb{Z}}$ consisting of the integral elements of $V_{\mathbb{R}}$.

The action of $G_{\mathbb{Z}}$ on $V_{\mathbb{R}}$ (or $V_{\mathbb{Z}}$) has a unique polynomial invariant. To see this, notice first that the action of $\mathrm{GL}_3(\mathbb{Z})$ on V has four independent polynomial invariants, namely the coefficients a, b, c, d of the binary cubic form

$$f(x, y) = f_{(A,B)}(x, y) = 4 \cdot \mathrm{Det}(Ax - By),$$

where $(A, B) \in V$. We call $f(x, y)$ the *binary cubic form invariant* of the element $(A, B) \in V$.

Next, $\mathrm{GL}_2(\mathbb{Z})$ acts on the binary cubic form $f(x, y)$, and it is well-known that this action has exactly one polynomial invariant, namely the discriminant $\mathrm{Disc}(f)$. Thus the unique polynomial invariant for the action of $G_{\mathbb{Z}}$ on $V_{\mathbb{Z}}$ is $\mathrm{Disc}(4 \cdot \mathrm{Det}(Ax - By))$. We call this fundamental invariant the *discriminant* $\mathrm{Disc}(A, B)$ of the pair (A, B). (The factor 4 is included to insure that any pair of integral ternary quadratic forms has integral discriminant.)

The orbits of $G_{\mathbb{Z}}$ on $V_{\mathbb{Z}}$ have an important arithmetic significance. Recall that a *quartic ring* is any ring that is isomorphic to $\mathbb{Z}^4$ as a $\mathbb{Z}$-module; for example, an order in a quartic number field is a quartic ring. In [3], we showed how quartic rings may be parametrized in terms of the $G_{\mathbb{Z}}$-orbits on $V_{\mathbb{Z}}$:

THEOREM 6. *There is a canonical bijection between the set of $G_{\mathbb{Z}}$-equivalence classes of elements $(A,B) \in V_{\mathbb{Z}}$ and the set of isomorphism classes of pairs (Q,R), where Q is a quartic ring and R is a cubic resolvent ring of Q. Under this bijection, we have* $\mathrm{Disc}(A,B) = \mathrm{Disc}(Q) = \mathrm{Disc}(R)$.

A cubic resolvent of a quartic ring Q is a cubic ring R equipped with a certain quadratic *resolvent mapping* $Q \to R$, whose precise definition will not be needed here (see [3] for details). In view of Theorem 6, it is natural to try to understand the number of $G_{\mathbb{Z}}$-orbits on $V_{\mathbb{Z}}$ having absolute discriminant at most X, as $X \to \infty$. The number of integral orbits on $V_{\mathbb{Z}}$ having a fixed discriminant D is called a "class number", and we wish to understand the behavior of this class number on average.

From the point of view of Theorem 6, we would like to restrict the elements of $V_{\mathbb{Z}}$ under consideration to those which are "irreducible" in an appropriate sense. More precisely, we call a pair (A,B) of integral ternary quadratic forms in $V_{\mathbb{Z}}$ *absolutely irreducible* if

- A and B do not possess a common zero as conics in $\mathbb{P}^2(\mathbb{Q})$; and
- the binary cubic form $f(x,y) = \mathrm{Det}(Ax - By)$ is irreducible over $\mathbb{Q}$.

Equivalently, (A,B) is absolutely irreducible if A and B possess a common zero in $\mathbb{P}^2$ having field of definition K, where K is a quartic number field whose Galois closure has Galois group either A_4 or S_4 over $\mathbb{Q}$. In terms of Theorem 6, absolutely irreducible elements in $V_{\mathbb{Z}}$ correspond to pairs (Q,R) where Q is an order in either an A_4 or S_4-quartic field. The main result of this section is the following theorem:

THEOREM 7. *Let $N(V_{\mathbb{Z}}^{(i)};X)$ denote the number of $G_{\mathbb{Z}}$-equivalence classes of absolutely irreducible elements $(A,B) \in V_{\mathbb{Z}}$ having $4-2i$ zeros in $\mathbb{P}^2(\mathbb{R})$ and satisfying $|\mathrm{Disc}(A,B)| < X$. Then*

$$\begin{aligned}
&\text{(a)} \quad \lim_{X\to\infty} \frac{N(V_{\mathbb{Z}}^{(0)};X)}{X} &&= \frac{\zeta(2)^2\zeta(3)}{48};\\
&\text{(b)} \quad \lim_{X\to\infty} \frac{N(V_{\mathbb{Z}}^{(1)};X)}{X} &&= \frac{\zeta(2)^2\zeta(3)}{8};\\
&\text{(c)} \quad \lim_{X\to\infty} \frac{N(V_{\mathbb{Z}}^{(2)};X)}{X} &&= \frac{\zeta(2)^2\zeta(3)}{16}.
\end{aligned}$$

Theorem 7 is proved in several steps. In Subsection 2.1, we outline the necessary reduction theory needed to establish some particularly useful fundamental domains for the action of $G_{\mathbb{Z}}$ on $V_{\mathbb{R}}$. In Subsection 2.2, we describe a new "averaging" method that allows one to efficiently count points in various components of these fundamental domains in terms of their volumes. In Subsections 2.3–2.5, we investigate the distribution of reducible and irreducible integral points within these fundamental domains. The volumes of the resulting "irreducible" components of these fundamental domains are then computed in the final Subsection 2.6, proving Theorem 7.

In Section 3, we will show how similar counting methods—together with a sieving process—can be used to prove Theorems 1–5.

2.1. *Reduction theory.* The action of $G_{\mathbb{R}} = \mathrm{GL}_2(\mathbb{R}) \times \mathrm{SL}_3(\mathbb{R})$ on $V_{\mathbb{R}}$ has three nondegenerate orbits $V_{\mathbb{R}}^{(0)}$, $V_{\mathbb{R}}^{(1)}$, $V_{\mathbb{R}}^{(2)}$, where $V_{\mathbb{R}}^{(i)}$ consists of those elements (A,B) in $V_{\mathbb{R}}$ having $4-2i$ common zeros in $\mathbb{P}^2(\mathbb{R})$. We wish to understand the number $N(V_{\mathbb{Z}}^{(i)};X)$ of absolutely irreducible $G_{\mathbb{Z}}$-orbits on $V_{\mathbb{Z}}^{(i)}$ having absolute discriminant less than X $(i=0,1,2)$. We accomplish this by counting the number of integer points of absolute discriminant less than X in suitable fundamental domains for the action of $G_{\mathbb{Z}}$ on $V_{\mathbb{R}}$.

These fundamental regions are constructed as follows. First, let $\mathcal{F}$ denote a fundamental domain for the action of $G_{\mathbb{Z}}$ on $G_{\mathbb{R}}$ by left multiplication. We may assume that $\mathcal{F} \subset G_{\mathbb{R}}$ is semi-algebraic and connected, and that it is contained in a standard Siegel set, i.e., $\mathcal{F} \subset N'A'K\Lambda$, where

$$\begin{aligned}
K &= \{\text{orthogonal transformations in } G_{\mathbb{R}}\};\\
A' &= \{a(t_1,t_2,t_3) : 0 < t_1^{-1} \le c_1\, t_1,\ 0 < (t_2t_3)^{-1} \le c_1\, t_2 \le c_1^2\, t_3\},\\
&\qquad \text{where } a(t_1,t_2,t_3) = \left(\begin{pmatrix} t_1^{-1} & \\ & t_1 \end{pmatrix}, \begin{pmatrix} (t_2t_3)^{-1} & & \\ & t_2 & \\ & & t_3 \end{pmatrix}\right); \quad \text{or}\\
A' &= \{a(s_1,s_2,s_3) : s_1 \ge 1/\sqrt{c_1},\ s_2,s_3 \ge 1/\sqrt[3]{c_1}\},\\
&\qquad \text{where } a(s_1,s_2,s_3) = \left(\begin{pmatrix} s_1^{-1} & \\ & s_1 \end{pmatrix}, \begin{pmatrix} s_2^{-2}s_3^{-1} & & \\ & s_2s_3^{-1} & \\ & & s_2s_3^2 \end{pmatrix}\right);\\
N' &= \{n(u_1,u_2,u_3,u_4) : |u_1|,|u_2|,|u_3|,|u_4| \le c_2\},\\
&\qquad \text{where } n(u_1,u_2,u_3,u_4) = \left(\begin{pmatrix} 1 & \\ u_1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & & \\ u_2 & 1 & \\ u_3 & u_4 & 1 \end{pmatrix}\right);\\
\Lambda &= \{\lambda : \lambda > 0\},\\
&\qquad \text{where } \lambda \text{ acts by } \left(\begin{pmatrix} \lambda & \\ & \lambda \end{pmatrix}, \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}\right),
\end{aligned}$$

and c_1, c_2 are absolute constants. For example, the well-known fundamental domains in $\mathrm{GL}_2(\mathbb{R})$ and $\mathrm{GL}_3(\mathbb{R})$ as constructed by Minkowski satisfy these conditions for $c_1 = 2/\sqrt{3}$ and $c_2 = 1/2$.

Next, for $i = 0, 1, 2$, let n_i denote the cardinality of the stabilizer in $G_{\mathbb{R}}$ of any element $v \in V_{\mathbb{R}}^{(i)}$. (One easily checks that $n_i = 24, 4, 8$ for $i = 0, 1, 2$ respectively.) Then for any $v \in V_{\mathbb{R}}^{(i)}$, $\mathcal{F}v$ will be the union of n_i fundamental domains for the action of $G_{\mathbb{Z}}$ on $V_{\mathbb{R}}^{(i)}$. Since this union is not necessarily disjoint, $\mathcal{F}v$ is best viewed as a multiset, where the multiplicity of a point x in $\mathcal{F}v$ is given by the cardinality of the set $\{g \in \mathcal{F} \mid gv = x\}$. Evidently, this multiplicity is a number between 1 and n_i.

Furthermore, since $\mathcal{F}v$ is a polynomial image of a semi-algebraic set $\mathcal{F}$, the theorem of Tarski and Seidenberg on quantifier elimination ([25], [24]) implies that $\mathcal{F}v$ is a semi-algebraic multiset in $V_{\mathbb{R}}$; here by a *semi-algebraic multiset* $\mathcal{R}$ we mean a multiset whose underlying subsets $\mathcal{R}_k$ of elements in $\mathcal{R}$ having multiplicity k are semi-algebraic for all $1 \leq k < \infty$. The semi-algebraicity of $\mathcal{F}v$ will play an important role in what follows (cf. Lemmas 9 and 15).

For any $v \in V_{\mathbb{R}}^{(i)}$, we have noted that the multiset $\mathcal{F}v$ is the union of n_i fundamental domains for the action of $G_{\mathbb{Z}}$ on $V_{\mathbb{R}}^{(i)}$. However, not all elements in $G_{\mathbb{Z}}\backslash V_{\mathbb{Z}}$ will be represented in $\mathcal{F}v$ exactly n_i times. In general, the number of times the $G_{\mathbb{Z}}$-equivalence class of an element $x \in V_{\mathbb{Z}}$ will occur in $\mathcal{F}v$ is given by $n_i/m(x)$, where $m(x)$ denotes the size of the stabilizer of x in $G_{\mathbb{Z}}$. Since we have shown in [3] that the stabilizer in $G_{\mathbb{Z}}$ of an *absolutely irreducible* element $(A, B) \in V_{\mathbb{Z}}$ is always trivial, we conclude that, for any $v \in V_{\mathbb{R}}^{(i)}$, the product $n_i \cdot N(V_{\mathbb{Z}}^{(i)}; X)$ is exactly equal to the number of absolutely irreducible integer points in $\mathcal{F}v$ having absolute discriminant less than X.

Thus to estimate $N(V_{\mathbb{Z}}^{(i)}; X)$, it suffices to count the number of integer points in $\mathcal{F}v$ for some $v \in V_{\mathbb{R}}^{(i)}$. The number of such integer points can be difficult to count in a single such $\mathcal{F}v$ (see e.g., [13], [2]), so instead we average over many $\mathcal{F}v$ by averaging over certain v lying in a box H.

2.2. *Averaging over fundamental domains.* Let $H = \{(A, B) \in V_{\mathbb{R}} : |a_{ij}|, |b_{ij}| \leq 10 \text{ for all } i, j;\ |\mathrm{Disc}(A, B)| \geq 1\}$, and let $\Phi = \Phi_H$ denote the characteristic function of H. Then since $\mathcal{F}v$ is the union of n_i fundamental domains for the action of $G_{\mathbb{Z}}$ on $V^{(i)} = V_{\mathbb{R}}^{(i)}$, we have

$$N(V_{\mathbb{Z}}^{(i)}; X) = \frac{\int_{v \in V^{(i)}} \Phi(v) \cdot \#\{x \in \mathcal{F}v \cap V_{\mathbb{Z}}^{(i)} \text{ abs. irr.} : 0 < |\mathrm{Disc}(x)| < X\}\ |\mathrm{Disc}(v)|^{-1} dv}{n_i \cdot \int_{v \in V^{(i)}} \Phi(v)\ |\mathrm{Disc}(v)|^{-1} dv}, \tag{4}$$

where points in $\mathcal{F}v \cap V_{\mathbb{Z}}^{(i)}$ are as usual counted according to their multiplicities in $\mathcal{F}v$. The denominator on the right-hand side of (4) is, by construction, a finite absolute constant M_i greater than zero. We have chosen to use the measure $|\mathrm{Disc}(v)|^{-1} dv$ because it is a $G_{\mathbb{R}}$-invariant measure.

More generally, for any $G_{\mathbb{Z}}$-invariant set $S \subset V_{\mathbb{Z}}$, we may speak of the number $N(S; X)$ of irreducible $G_{\mathbb{Z}}$-orbits on S having absolute discriminant less than X. Then $N(S; X)$ can be expressed similarly as

(5)

$$N(S;X) = \sum_{i=0}^{2} \frac{\int_{v\in V^{(i)}} \Phi(v)\cdot \#\{x\in\mathcal{F}v\cap S \text{ abs. irr.} : 0<|\mathrm{Disc}(x)|<X\}\, |\mathrm{Disc}(v)|^{-1}dv}{n_i\cdot \int_{v\in V^{(i)}}\Phi(v)\,|\mathrm{Disc}(v)|^{-1}dv}.$$

We shall use this definition of $N(S;X)$ for any $S\subset V_{\mathbb{Z}}$, even if S is not $G_{\mathbb{Z}}$-invariant. Note that for disjoint $S_1,S_2\subset V_{\mathbb{Z}}$, we have $N(S_1\cup S_2) = N(S_1)+N(S_2)$.

Using the fact that $|\mathrm{Disc}(v)|^{-1}dv$ is the unique $G_{\mathbb{R}}$-invariant measure on $V^{(i)}$ (up to scaling), we may also express formula (5) for $N(S;X)$ as an integral over $\mathcal{F}^{-1}\subset G_{\mathbb{R}}$. Let dg be a left-invariant Haar measure on $G_{\mathbb{R}}$, which is uniquely defined up to scaling. Then we may write

$$(6)\qquad \begin{aligned} N(S;X) &= \sum_{i=0}^{2}\frac{1}{M_i}\int_{v\in V^{(i)}} \sum_{\substack{x\in\mathcal{F}v\cap S \text{ abs. irr.}\\ |\mathrm{Disc}(x)|<X}} \Phi(v)\ |\mathrm{Disc}(v)|^{-1}dv\\ &= \sum_{i=0}^{2}\frac{c'}{M_i}\int_{g\in\mathcal{F}^{-1}} \sum_{\substack{x\in V^{(i)}\cap S \text{ abs. irr.}\\ |\mathrm{Disc}(x)|<X}} \Phi(gx)\,dg,\end{aligned}$$

where c' is an absolute constant depending only on the scaling of the Haar measure dg. In particular, since $\mathcal{F}^{-1}\subset KA'^{-1}N'\Lambda\subset KN'A'^{-1}\Lambda$, we have the upper bound

$$(7)\quad N(S;X)\ll \int_{g\in KN'A'^{-1}\Lambda} \sum_{\substack{x\in S \text{ abs. irr.}\\ |\mathrm{Disc}(x)|<X}} \Phi(kna^{-1}\lambda x)\ s_1^{-2}s_2^{-6}s_3^{-6}\ d^\times\lambda\, d^\times s\, dn\, dk.$$

Note that, in the latter integral, it suffices to restrict $\lambda\in\Lambda$ to within the range $[X^{-1/12},c]$, where $c=(\max\{|\mathrm{Disc}(x)| : x\in H\})^{1/12}$ is an absolute constant. Indeed, if $x\in S$ with $1\le|\mathrm{Disc}(x)|<X$ and λ is outside the range $[X^{-1/12},c]$, then $|\mathrm{Disc}(kna^{-1}\lambda x)| = \lambda^{12}|\mathrm{Disc}(x)|$ will lie outside the range $[1,c^{12}]$; in that case, $kna^{-1}\lambda x\notin H$ and the integrand will be zero.

Now since K and N' are compact, there exists a compact set H' such that $H'\supset N'KH$. In fact, we may set

$$H' = \{(A,B)\in V_{\mathbb{R}} : |a_{ij}|,|b_{ij}|\le 60 \text{ for all } i,j;\ |\mathrm{Disc}(A,B)|\ge 1\}$$

as it is easy to check that the latter set contains $N'KH$. Let Ψ denote the characteristic function of H'. Then (7) implies

$$(8)\qquad N(S;X)\ll \int_{\lambda=X^{-\frac{1}{12}}}^{c}\int_{s_1,s_2,s_3=\frac12}^{\infty} \sigma(S)\ \ s_1^{-2}s_2^{-6}s_3^{-6}d^\times s\, d^\times\lambda,$$

where $\sigma(S)=\sigma(S;\lambda,s_1,s_2,s_3)$ is given by

$$\sigma(S) = \sum_{\substack{(A,B)\in S \text{ abs. irr.}\\ |\mathrm{Disc}(A,B)|<X}} \Psi\left(\lambda\cdot a(s_1,s_2,s_3)^{-1}(A,B)\right).$$

Noting that $2\,\lambda\cdot a(s_1,s_2,s_3)^{-1}(A,B)$ is

(9)

$$\left(\begin{bmatrix} 2\,\lambda s_1 s_2^4 s_3^2 a_{11} & \lambda s_1 s_2 s_3^2 a_{12} & \lambda s_1 s_2 s_3^{-1} a_{13} \\ \lambda s_1 s_2 s_3^2 a_{12} & 2\,\lambda s_1 s_2^{-2} s_3^2 a_{22} & \lambda s_1 s_2^{-2} s_3^{-1} a_{23} \\ \lambda s_1 s_2 s_3^{-1} a_{13} & \lambda s_1 s_2^{-2} s_3^{-1} a_{23} & 2\,\lambda s_1 s_2^{-2} s_3^{-4} a_{33} \end{bmatrix}, \begin{bmatrix} 2\,\lambda s_1^{-1} s_2^4 s_3^2 a_{11} & \lambda s_1^{-1} s_2 s_3^2 a_{12} & \lambda s_1^{-1} s_2 s_3^{-1} a_{13} \\ \lambda s_1^{-1} s_2 s_3^2 a_{12} & 2\,\lambda s_1^{-1} s_2^{-2} s_3^2 a_{22} & \lambda s_1^{-1} s_2^{-2} s_3^{-1} a_{23} \\ \lambda s_1^{-1} s_2 s_3^{-1} a_{13} & \lambda s_1^{-1} s_2^{-2} s_3^{-1} a_{23} & 2\,\lambda s_1^{-1} s_2^{-2} s_3^{-4} a_{33} \end{bmatrix}\right),$$

we see that $\lambda\cdot a(s_1,s_2,s_3)^{-1}(A,B)$ will lie in H' only if (A,B) lies in the box defined by the inequalities

(10)

$$|a_{11}|\le\frac{60}{\lambda s_1 s_2^4 s_3^2};\ |a_{12}|\le\frac{60}{\lambda s_1 s_2 s_3^2};\ |a_{13}|\le\frac{60 s_3}{\lambda s_1 s_2};\ \ |a_{22}|\le\frac{60 s_2^2}{\lambda s_1 s_3^2};\ \ |a_{23}|\le\frac{60 s_2^2 s_3}{\lambda s_1};\ \ \ |a_{33}|\le\frac{60 s_2^2 s_3^4}{\lambda s_1};$$

$$|b_{11}|\le\frac{60 s_1}{\lambda s_2^4 s_3^2};\ \ \ |b_{12}|\le\frac{60 s_1}{\lambda s_2 s_3^2};\ \ \ |b_{13}|\le\frac{60 s_1 s_3}{\lambda s_2};\ |b_{22}|\le\frac{60 s_1 s_2^2}{\lambda s_3^2};\ |b_{23}|\le\frac{60 s_1 s_2^2 s_3}{\lambda};\ |b_{33}|\le\frac{60 s_1 s_2^2 s_3^4}{\lambda}.$$

Hence $\sigma(S)$ is at most the number of absolutely irreducible points in S lying in the box (10). In practice, we will choose our sets $S\subset V_{\mathbb{Z}}$ for which it is easy to estimate the number of points in S lying in the box (10). This will allow for accurate estimates of $N(S;X)$.

We note that the same counting method may be used even if we are interested in counting both reducible and irreducible orbits in $V_{\mathbb{Z}}$. For any set $S\subset V_{\mathbb{Z}}$, let $N^*(S;X)$ be defined by (5), but where the phrase "abs. irr." is removed. Thus for a $G_{\mathbb{Z}}$-invariant set $S\subset V_{\mathbb{Z}}$, $N^*(S;X)$ counts the total number of $G_{\mathbb{Z}}$-orbits in S having absolute discriminant nonzero and less than X (not just the irreducible ones). By the same reasoning, we have

$$N^*(S;X)\ll\int_{\lambda=X^{-\frac{1}{12}}}^{c}\int_{s_1,s_2,s_3=\frac{1}{2}}^{\infty}\sigma^*(S)\ \ s_1^{-2}s_2^{-6}s_3^{-6}d^\times s\,d^\times\lambda, \tag{11}$$

where $\sigma^*(S)=\sigma(S;\lambda,s_1,s_2,s_3)$ denotes the number of integer points in S satisfying (10).

The expression (5) for $N(S;X)$, its analogue for $N^*(S,X)$, the upper bounds (8) and (11), and the inequalities (10) will be useful in the sections that follow.

2.3. *Preliminary estimates.* We begin with some estimates that must be satisfied by the coefficients of any element $(A,B)\in\mathcal{F}v$, where $v\in H$.

LEMMA 8. *Let $v\in H$. Suppose $(A,B)\in\mathcal{F}v$ has entries given by* (3) *and satisfies $|\mathrm{Disc}(A,B)|<X$. Let S be any multiset consisting of elements of the form a_{ij} or b_{ij}. Let m denote the number of a's which occur in S, and let $n=|S|-m$ denote the number of b's; let i, j, and $k=2\,|S|-i-j$ denote the number of indices in S equal to* 1, 2, *and* 3 *respectively. If $m\ge n$, $2i\ge j+k$, and $i+j\ge 2k$, then*

$$\prod_{s\in S}s=O(X^{|S|/12}).$$

Proof. Note that $\mathcal{F}v \subset \Lambda' N' A' K v$, where N', A', and K are as in Section 2.1 and $\Lambda' = \{\lambda \in \mathbb{R} : 0 < \lambda < X^{1/12}\}$. For a multiset S as in the lemma, it is clear that the value of $f = \prod_{s\in S} s$ is bounded on Kv, since K and H are compact. Next, the values of f on $A'Kv$ are simply $s_1^{n-m} s_2^{j+k-2i} s_3^{2k-i-j}$ times the values of f on Kv. If $m \geq n$, $2i \geq j+k$, and $i+j \geq 2k$, then it is clear that $s_1^{n-m} s_2^{j+k-2i} s_3^{2k-i-j}$ is absolutely bounded, and hence the values of f on $A'Kv$ are also bounded. Finally, N' is compact, and it acts only by lower triangular transformations; thus f also takes bounded values on $N'A'Kv$. Therefore, the values of f on $\Lambda' N' A' K v$ are at most $O(X^{|S|/12})$ in size. This is the desired conclusion. $\square$

Lemma 8 gives those estimates on the entries of (A, B) that follow immediately from the fact that $\mathcal{F}$ is contained in a Siegel set.

The following two lemmas will also be useful. The first is essentially due to Davenport [12], [14]. To state the lemma, we require the following simple definitions. A multiset $\mathcal{R} \subset \mathbb{R}^n$ is said to be *measurable* if $\mathcal{R}_k$ is measurable for all k, where $\mathcal{R}_k$ denotes the set of those points in $\mathcal{R}$ having a fixed multiplicity k. Given a measurable multiset $\mathcal{R} \subset \mathbb{R}^n$, we define its volume in the natural way; that is, $\mathrm{Vol}(\mathcal{R}) = \sum_k k \cdot \mathrm{Vol}(\mathcal{R}_k)$, where $\mathrm{Vol}(\mathcal{R}_k)$ denotes the usual Euclidean volume of $\mathcal{R}_k$.

LEMMA 9. *Let $\mathcal{R}$ be a bounded, semi-algebraic multiset in $\mathbb{R}^n$ having maximum multiplicity m, where $\mathcal{R}$ is defined by at most k polynomial inequalities each having degree at most ℓ. Then the number of integer lattice points (counted with multiplicity) contained in the region $\mathcal{R}$ is*

$$\mathrm{Vol}(\mathcal{R}) + O(\max\{\mathrm{Vol}(\bar{\mathcal{R}}), 1\}),$$

where $\mathrm{Vol}(\bar{\mathcal{R}})$ denotes the greatest d-dimensional volume of any projection of $\mathcal{R}$ onto a coordinate subspace obtained by equating $n-d$ coordinates to zero, where d takes all values from 1 to $n-1$. The implied constant in the second summand depends only on n, m, k, and ℓ.

Although Davenport states Lemma 9 only for compact semi-algebraic sets, his proof adapts without essential change to the more general case of bounded semi-algebraic multisets.

The following effective special case of Lemma 9 will be particularly useful.

LEMMA 10. *Let $c > 0$, and let $\mathcal{B}$ be a closed box in $\mathbb{R}^n$ each of whose faces is parallel to a coordinate hyperplane and each of whose edges has length at least c. Then the number of integer points in $\mathcal{B}$ is at most $C \cdot \mathrm{Vol}(\mathcal{B})$, where C is an absolute constant depending only on c.*

The proof of Lemma 10 is trivial. Furthermore, it is easy to see that we may take $C = \max\{\lceil c\rceil/c,\, 1+1/\lceil c\rceil\}^n$, with equality if and only if $\mathcal{B}$ is an appropriately placed n-dimensional hypercube in $\mathbb{R}^n$ whose edges each have length either c or $\lceil c\rceil$ (whichever gives the bigger value of C).

Notation. In what follows, we use ϵ to denote any positive real number. Thus we say "$f(X) = O(X^{1+\epsilon})$" if $f(X) = O(X^{1+\epsilon})$ for any $\epsilon > 0$.

2.4. *Estimates on reducible pairs* (A, B). In this section we describe the relative frequencies with which absolutely irreducible elements sit inside various parts of the multiset $\mathcal{F}v$, as v varies over the box H.

LEMMA 11. *Let v take a random value in H uniformly with respect to the measure $|\mathrm{Disc}(v)|^{-1}\,dv$. Then the expected number of absolutely irreducible elements $(A, B) \in \mathcal{F}v \cap V_{\mathbb{Z}}$ such that $a_{11} = 0$ and $|\mathrm{Disc}(A, B)| < X$ is $O(X^{11/12})$.*

Proof. Let $V(0)$ denote the set of $(A, B) \in V_{\mathbb{R}}$ such that $a_{11} = 0$. Note that if an element $(A, B) \in V(0)$ is absolutely irreducible, then we must have $b_{11} \neq 0$, for otherwise $(1, 0, 0) \in \mathbb{P}^2(\mathbb{Q})$ would be a common zero of A and B.

We wish to show that $N(V(0); X)$, as defined by (5), is $O(X^{11/12})$. To estimate $N(V(0); X)$, we partition $V(0)$ into two sets: $V(0*)$, consisting of those elements $(A, B) \in V(0)$ for which $a_{12} \neq 0$; and $V(00)$, consisting of those (A, B) where both $a_{11} = a_{12} = 0$. Then we have $N(V(0); X) = N(V(0*); X) + N(V(00); X)$. We estimate the latter two terms in two cases.

Case I. $N(V(0*); X)$. In this case, estimate (8) becomes

$$N(V(0*); X) \ll \int_{\lambda = X^{-\frac{1}{12}}}^{c} \int_{s_1, s_2, s_3 = \frac{1}{2}}^{\infty} \sigma(V(0*))\, s_1^{-2} s_2^{-6} s_3^{-6} d^\times s\, d^\times\lambda, \tag{12}$$

where $\sigma(V(0*))$ is at most the number of integer points in the box defined by the inequalities (10) together with the conditions

$$a_{11} = 0,\ |a_{12}| \geq 1,\ |b_{11}| \geq 1. \tag{13}$$

The number of integer points $(a_{12}, \cdots, b_{33}) \in \mathbb{R}^{11}$ satisfying the latter requirements can be positive only if the quantities $\frac{60}{\lambda s_1 s_2 s_3^2}$ and $\frac{60 s_1}{\lambda s_2^4 s_3^2}$ are each at least 1, since $|a_{12}|$, $|b_{11}| \geq 1$. In that case, the conditions (10) and (13) define a union of four boxes in $\mathbb{R}^{11}$, each of whose sidelengths is seen to be bounded below by 2^{-11}. By Lemma 10, it follows that the number of integer points in $\mathcal{B}$ is bounded above by an absolute constant times $\mathrm{Vol}(\mathcal{B})$. Since $\mathrm{Vol}(\mathcal{B}) \ll \lambda^{-11} s_1 s_2^4 s_3^2$, we have

$$\sigma(V(0*)) \ll \lambda^{-11} s_1 s_2^4 s_3^2. \tag{14}$$

Equation (12) then implies

$$(15)\quad N(V(0*);X) \ll \int_{\lambda=X^{-\frac{1}{12}}}^{c} \int_{s_1,s_2,s_3=\frac{1}{2}}^{\infty} \lambda^{-11} s_1^{-1} s_2^{-2} s_3^{-4} d^\times s\, d^\times\lambda = O(X^{11/12})$$

as desired.

Case II. $N(V(00);X)$. If we have (A,B) with $a_{11}=a_{12}=0$ then $a_{13}\neq 0$ and $a_{22}\neq 0$, or else the cubic form invariant $f(x,y)=\mathrm{Det}(Ax-By)$ would be reducible. Therefore, by estimate (8), the expected number of absolutely irreducible elements $(A,B)\in V(00)$ with $|\mathrm{Disc}(A,B)|<X$ is

$$(16)\qquad N(V(00);X) \ll \int_{\lambda=X^{-\frac{1}{12}}}^{c} \int_{s_1,s_2,s_3=\frac{1}{2}}^{\infty} \sigma(V(00))\, s_1^{-2} s_2^{-6} s_3^{-6} d^\times s\, d^\times\lambda,$$

where $\sigma(V(0*))$ is bounded above by the number of integer points in the box defined by the inequalities (10) and the conditions

$$(17)\qquad a_{11}=0,\ a_{12}=0,\ |a_{13}|\geq 1,\ |a_{22}|\geq 1,\ |b_{11}|\geq 1.$$

The conditions (10) and (17) define a region $\mathcal{B}\subset\mathbb{R}^{10}$. This region can have an integer point only if the quantities $\frac{60s_3}{\lambda s_1 s_2}$, $\frac{60s_2^2}{\lambda s_1 s_3^2}$, and $\frac{60 s_1}{\lambda s_2^4 s_3^2}$ are each at least 1. In that case, we observe that $\mathcal{B}$ is the union of eight boxes each of whose sidelengths is at least 2^{-8}. By Lemma 10, the number of integer points in $\mathcal{B}$ is at most $C(2^{-8})\cdot\mathrm{Vol}(\mathcal{B})\ll \lambda^{-10}s_1^2s_2^5s_3^4$. Hence from (16) we have

$$(18)\quad N(V(00);X) \ll \int_{\lambda=X^{-\frac{1}{12}}}^{c} \int_{s_1,s_2,s_3} \lambda^{-10} s_2^{-1} s_3^{-2} d^\times s\, d^\times\lambda = O(X^{10/12}\log X),$$

since equations (10) and (17) together imply that $s_1\leq\frac{60}{\lambda}\leq 60X^{1/12}$. This yields the lemma. □

Thus, for the purposes of proving Theorem 7, we may assume that $a_{11}\neq 0$.

LEMMA 12. *Let $v\in H$. The number of $(A,B)\in\mathcal{F}v$ such that $a_{11}\neq 0$, $|\mathrm{Disc}(A,B)|<X$, and $f(x,y)=\mathrm{Det}(Ax-By)$ is reducible is $O(X^{11/12})$.*

Proof. Any cubic ring $R=R(f)$ of discriminant n such that $f(x,y)$ is a reducible cubic form sits in a unique cubic $\mathbb{Q}$-algebra $K=R\otimes\mathbb{Q}\cong\mathbb{Q}\oplus F$, where F is a certain quadratic $\mathbb{Q}$-algebra (indeed, F depends only on the squarefree part of n). Let us write $\mathrm{Disc}(R)=k^2\mathrm{Disc}(K)$. Then the number of quartic $\mathbb{Q}$-algebras L having discriminant dividing $\mathrm{Disc}(R)=k^2\mathrm{Disc}(K)$, and such that the cubic resolvent of L is K, is $O(h_2^*(K)\mathrm{Disc}(R)^\epsilon)$ by the work of Baily [1].[§] Since K is of the form $\mathbb{Q}\oplus F$, where F is a quadratic $\mathbb{Q}$-algebra,

[§]Although Baily states all results for "cubic fields", it is clear that his arguments hold also when every occurrence of "field" is replaced by "étale $\mathbb{Q}$-algebra".

we have $h_2^*(K) = O(\mathrm{Disc}(K)^\epsilon)$ by genus theory. Hence the total number of possibilities for the quartic $\mathbb{Q}$-algebra L, given $R = R(f)$, is $O(\mathrm{Disc}(R)^\epsilon)$.

Now any quartic ring Q such that the cubic resolvent ring of Q is R must be an order in such an L, and the index of this order in $\mathcal{O}_L$ (the ring of integers of L) must divide k. In particular, for a fixed choice of L the number of $Q \subseteq L$ with $R^{\mathrm{res}}(Q) = R(f)$ is at most the number of orders of index k in $\mathcal{O}_L$. For any integer $k > 0$, let $\mathrm{EP}(n)$ denote the product of all factors p^e occurring in the prime power decomposition of n such that $e \geq 8$. Then it follows from a result of Nakagawa [20, Theorem 1] that the number of orders of index k in the ring of integers in an étale quartic $\mathbb{Q}$-algebra L is at most $O(\mathrm{EP}(k^2)^{1/4+\epsilon})$, independent of L.

Let $s = 16/27$. We divide the set S of reducible cubic forms $f(x, y)$ into two sets: S_1, the set of all reducible cubic forms f with $\mathrm{EP}(\mathrm{Disc}(f)) \geq \mathrm{Disc}(f)^s$, and S_2, the set of all reducible cubic forms f with $\mathrm{EP}(\mathrm{Disc}(f)) < \mathrm{Disc}(f)^s$.

We treat first the $(A, B) \in \mathcal{F}v$ with $f(x, y) \in S_1$ and $|\mathrm{Disc}(f)| < X$. It is a standard fact that the number of positive integers $n < X$ such that $\mathrm{EP}(n) \geq n^s$ is $O(X^{1-\frac{7}{8}s+\epsilon})$. Furthermore, it is easy to see (see e.g., Datskovsky-Wright [11], Nakagawa [21]) that the number of orders of a given index k in the maximal order of a cubic $\mathbb{Q}$-algebra K is at most $O(k^{1/3+\epsilon})$, independent of K; it follows that the number of reducible $f(x, y)$ with a given discriminant n is at most $O(n^{1/6+\epsilon})$. Hence the total number of classes of reducible cubic forms $f \in S_1$ satisfying $0 < |\mathrm{Disc}(f)| < X$ is at most $O(X^{1-\frac{7}{8}s+\epsilon} \cdot X^{\frac{1}{6}+\epsilon})$.

Finally, given an $f \in S_1$ with $0 < |\mathrm{Disc}(f)| < X$, the number of quartic $\mathbb{Q}$-algebras L of discriminant at most $\mathrm{Disc}(R(f))$, such that the cubic resolvent of L is $K = R(f) \otimes \mathbb{Q}$, is $O(\mathrm{Disc}(f)^\epsilon) = O(X^\epsilon)$; and the maximal number of orders Q of index k in $\mathcal{O}_L$ is at most $O(\mathrm{EP}(k^2)^{1/4+\epsilon}) = O(X^{1/4+\epsilon})$. We conclude that the total number of $(A, B) \in \mathcal{F}v$ with $f(x, y) \in S_1$ and $|\mathrm{Disc}(f)| < X$ is

$$O(X^{1-\frac{7}{8}s+\epsilon} \cdot X^{\frac{1}{6}+\epsilon} \cdot X^\epsilon \cdot X^{\frac{1}{4}+\epsilon}). \tag{19}$$

To similarly treat the $(A, B) \in \mathcal{F}v$ with $f(x, y) \in S_2$ and $|\mathrm{Disc}(f)| < X$, we may invoke a result of Davenport [13, Lemma 3], the proof of which implies that the total number of reducible forms $f(x, y) = ax^3 + bx^2y + cxy^2 + dy^3$ arising from an $(A, B) \in \mathcal{F}v$ such that $a \neq 0$ and $|\mathrm{Disc}(f)| < X$ is at most $O(X^{3/4+\epsilon})$. In particular, the total number of such cubic forms $f \in S_2$ is at most $O(X^{3/4+\epsilon})$. Now given an $f \in S_2$, the number of quartic $\mathbb{Q}$-algebras L having discriminant at most $\mathrm{Disc}(R(f))$, such that the cubic resolvent of L is $K = R(f) \otimes \mathbb{Q}$, is $O(\mathrm{Disc}(f)^\epsilon) = O(X^\epsilon)$; and the number of orders Q of index k in $\mathcal{O}_L$ is at most $O(\mathrm{EP}(k^2)^{1/4+\epsilon}) = O(k^{\frac{1}{2}s+\epsilon}) = O(X^{\frac{1}{4}s+\epsilon})$. Therefore, the total number of $(A, B) \in \mathcal{F}v$ with $f(x, y) \in S_2$, $a \neq 0$, and $|\mathrm{Disc}(f)| < X$ is

$$O(X^{\frac{3}{4}+\epsilon} \cdot X^\epsilon \cdot X^{\frac{1}{4}s+\epsilon}). \tag{20}$$

Choosing $s = 16/27$ yields $O(X^{97/108+\epsilon})$ in both (19) and (20), and thus both are $O(X^{11/12})$.

It remains only to show that the number of (A, B) satisfying the conditions of the lemma, for which $a = \mathrm{Det}(A) = 0$, is also at most $O(X^{11/12})$. To this end, note that $\mathrm{Det}(A) = 0$ is a quadratic equation in a_{23}, with nonzero leading coefficient a_{11}. It follows that once all entries of A except for a_{23} are fixed, then a_{23} too is determined up to at most two possibilities by the equation $\mathrm{Det}(A) = 0$.

Let S denote the set of all $(A, B) \in V_{\mathbb{Z}}$ such that $\mathrm{Det}(A) = 0$, so that the entry a_{23} of A is determined up to two possibilities by the other entries of A. Then estimate (11) applies to $N^*(S; X)$, where $\sigma^*(S)$ is the number of points in S in the region defined by (10) but where we assume a_{23} takes values in a set of cardinality at most two. Thus we may consider the 11-dimensional region $\mathcal{B}$ defined by (10) in the 11 variable entries of (A, B) excluding a_{23}. This region $\mathcal{B}$ can have an integer point only if $\frac{60}{\lambda s_1 s_2^4 s_3^2} \geq 1$ (since $|a_{11}|$ must be at least 1). In that case, $\mathcal{B}$ is seen to be a union of two boxes in $\mathbb{R}^{11}$ each of whose sidelengths is at least 2^{-14}; by Lemma 10, we have

$$\sigma^*(S) \ll 2 \cdot \mathrm{Vol}(\mathcal{B}) \ll 2 \cdot \lambda^{-11} s_1 s_2^{-2} s_3^{-1}$$

so that

$$N^*(S; X) \ll 2 \int_{\lambda = X^{-\frac{1}{12}}}^{c} \int_{s_1, s_2, s_3 = \frac{1}{2}}^{\infty} \lambda^{-11} s_1^{-1} s_2^{-8} s_3^{-7} \, d^\times s \, d^\times \lambda = O(X^{11/12}),$$

as was desired. □

Let T denote the set of twelve variables $\{a_{ij}, b_{ij}\}$. Note that $a_{11} \neq 0$ together with the estimate $a_{11}^2 t = O(X^{1/3})$ for $t \in T$ (Lemma 8) shows that

$$t = O(X^{1/3})$$

for all $t \in T$.

Lemma 13. *Let v take a random value in H uniformly with respect to the measure $|\mathrm{Disc}(v)|^{-1} \, dv$. Then the expected number of integer points $(A, B) \in \mathcal{F}v$ such that $a_{11} \neq 0$, $|\mathrm{Disc}(A, B)| < X$, and A and B have a common zero in $\mathbb{P}^2(\mathbb{Q})$ is $O(X^{11/12+\epsilon})$.*

Proof. We introduce some simple notation that will be needed during the course of the proof. First, let $R_1(y, z)$, $R_2(x, z)$, $R_3(x, y)$ denote the resultants of the two quadratic forms $A(x, y, z)$ and $B(x, y, z)$ with respect to the variables x, y, z respectively. The R_i's are thus binary quartic forms.

Next, denote by $A_{12}(x, y)$, $A_{13}(x, z)$, $A_{23}(y, z)$ the binary quadratic forms obtained from $A(x, y, z)$ by setting z, y, x equal to zero respectively. Define $B_{12}(x, y)$, $B_{13}(x, z)$, and $B_{23}(y, z)$ analogously. Associate with these pairs (A_{12}, B_{12}), (A_{13}, B_{13}), (A_{23}, B_{23}) of binary quadratic forms their discriminant invariants D_{12}, D_{13}, D_{23} given by

$$D_{ij} = \mathrm{Disc}(\mathrm{Det}(A_{ij} x - B_{ij} y)).$$

Equivalently, D_{ij} is the resultant of the binary quadratic forms $A_{ij}(x,y)$ and $B_{ij}(x,y)$ with respect to y, divided by x^4. The discriminants D_{ij} are forms of degree four in the entries of (A,B). We note also that D_{12} is the coefficient of x^4 in $R_2(x,z)$ and of y^4 in $R_1(y,z)$, with the analogous interpretations for D_{13} and D_{23}.

Now fix $v \in H$, and let (A,B) be an element in $\mathcal{F}v$ with $|\mathrm{Disc}(A,B)| < X$ for which A and B have a common rational zero $(r,s,t) \in \mathbb{P}^2(\mathbb{Q})$. We choose r,s,t to be integers having no common factor. If there is more than one rational zero, we choose (r,s,t) so that as many of the r,s,t are zero as possible. We write $r = (r,s)(r,t)r_0$, $s = (r,s)(s,t)s_0$, $t = (r,t)(s,t)t_0$, where (m,n) denotes the greatest common divisor of m and n (set $(m,0) = (0,n) = 1$ for convenience).

Let us consider first the case where $rst \neq 0$ (so that A and B have no common rational point in $\mathbb{P}^2$ with a coordinate equal to zero). To bound the number of possibilities for (A,B) in this case, we examine the discriminants D_{12}, D_{13}, D_{23}.

If any of these discriminants, say D_{12}, is equal to zero, then the corresponding pair of quadratic forms (A_{12},B_{12}) must have a common zero (r',s') in $\mathbb{P}^1$. By assumption, this zero cannot be rational, for otherwise $(r',s',0)$ would be a common rational zero of (A,B) having a zero coordinate. Therefore, if $D_{12}=0$, then A_{12}, B_{12} possess the same pair of conjugate zeros (defined over some quadratic extension of $\mathbb{Q}$), and thus A_{12} and B_{12} are scalar multiples of each other. Pick $u,v \in \mathbb{Z}$ such that $uA_{12} - vB_{12} = 0$. Then clearly $f(u,v) = \mathrm{Det}(uA - vB) = 0$, so that $f(x,y)$ is reducible over $\mathbb{Q}$. Such elements (A,B) with $f(x,y)$ reducible have already been handled, by Lemma 12.

We may therefore assume that $D_{12} \neq 0$, $D_{13} \neq 0$, and $D_{23} \neq 0$. If all a_{ij}, b_{ij} aside from possibly b_{23} are nonzero, then the estimate (Lemma 8)

$$\prod_{t \in T \setminus \{b_{23}\}} t = O(X^{11/12}) \tag{21}$$

implies that the number of nonzero choices for the variables in $T \setminus \{b_{23}\}$ is $O(X^{11/12+\epsilon})$. If some elements of $T \setminus \{b_{23}\}$ are equal to 0, we may replace those variables in (21) by a_{11}, and the estimate still remains true by Lemma 8. Thus the number of choices for the remaining nonzero variables in T is still $O(X^{11/12+\epsilon})$.

Once the variables in $T \setminus \{b_{23}\}$ have been chosen, they also determine the quantities D_{12} and D_{13}, which by assumption are nonzero. Since the coefficients of x^4 in $R_3(x,y)$ and $R_2(x,z)$ are D_{12} and D_{13} respectively, and $R_3(r,s) = R_2(r,t) = 0$, it follows that t_0 and s_0 divide D_{12} and D_{13} respectively. Thus the number of possibilities for s_0 and t_0 are bounded by the number of factors of D_{12} and D_{13} respectively. Since $D_{12}D_{13} = O(X^{2/3})$ by Lemma 8, the number of possibilities for s_0, t_0 is at most $O(X^{\epsilon})$. Now r divides (the nonzero quantity) $A_{23}(s,t)$, and as $A_{23}(s,t)$ is clearly at most $O(X^2)$ in absolute value, the number of choices for r is also at most $O(X^\epsilon)$. The factors (r,s), (r,t), and (s,t) are also determined up to $O(X^\epsilon)$ choices, as they are factors of r, r, and a_{11} respectively. Finally, since $B(r,s,t)=0$, the value of b_{23} is uniquely determined by $T\setminus\{b_{23}\}$, r, s, and t. Hence the number of choices for b_{23}, given $T\setminus\{b_{23}\}$, is at most $O(X^\epsilon)$, and so the total number of choices for T is $O(X^{11/12+\epsilon})$.

We consider next the cases where exactly one of r,s,t is equal to zero (so that A and B do not have a common rational point in $\mathbb{P}^2$ with two coordinates equal to zero).

If $r=0$ and $st\neq 0$, then

$$A_{23}(s,t)=B_{23}(s,t)=0. \tag{22}$$

We can assume that at least one of a_{22}, b_{22} (say b_{22}) and at least one of a_{33}, b_{33} (say b_{33}) is nonzero, for otherwise $(0,1,0)$ or $(0,0,1)$ would be a rational zero of (A,B) with two zero coordinates. Since

$$\prod_{t\in T\setminus\{a_{23},b_{23}\}} t = O(X^{10/12}) \tag{23}$$

(where as before zero variables are replaced by a_{11}), we see that the number of choices for $T\setminus\{a_{23},b_{23}\}$ is bounded by $O(X^{10/12+\epsilon})$. Once these choices are made, (22) implies that s divides b_{33} and t divides b_{22}; hence the number of possibilities for s and t is bounded by the number of factors of b_{33} and b_{22} respectively; so s and t can take at most $O(X^\epsilon)$ values (since b_{22} and b_{33} are both $O(X^{1/3})$). The values of a_{23} and b_{23} are then determined by $T\setminus\{a_{23},b_{23}\}$, r, s, and t. Thus the total number of possibilities for (A,B) in this case is $O(X^{10/12+\epsilon})$.

The case $s=0$, $rt\neq 0$ is handled similarly; the equation (23) is simply changed to

$$a_{11}\prod_{t\in T\setminus\{a_{13},b_{13}\}} t = O(X^{11/12}), \tag{24}$$

and we find in conclusion that there are at most $O(X^{11/12+\epsilon})$ choices for (A,B) in this case.

The case $t=0$, $rs\neq 0$ is a bit more difficult. Proceeding in the same manner, we find a_{12} and b_{12} are determined up to $O(X^\epsilon)$ possibilities once a_{11}, a_{22}, b_{11}, and b_{22} are fixed. However, equation (24) now becomes

$$a_{11}^2\prod_{t\in T\setminus\{a_{12},b_{12}\}} t = O(X^{12/12}), \tag{25}$$

and this does not yield a satisfactory estimate. Nevertheless, we can still show that the expected number of possibilities in this case, as v ranges over H, is at most $O(X^{10/12+\epsilon})$.

Indeed, let S denote the set of $(A,B)\in V_{\mathbb{Z}}$ such that A and B have a common zero of the form $(r,s,0)$ with $rs\neq 0$, so that a_{12} and b_{12} are determined up to $O(X^{\epsilon})$ possibilities by the remaining variables. Then estimate (11) applies to $N^*(S;X)$, where $\sigma^*(S)$ is the number of points in S in the region defined by (10) but where we assume a_{12},b_{12} take values in sets of cardinality at most $O(X^{\epsilon})$. Thus we may consider the 10-dimensional region $\mathcal{B}$ defined by (10) in the ten variables of $T\setminus\{a_{12},b_{12}\}$. This region $\mathcal{B}$ can have an integer point only if $\frac{60}{\lambda s_1 s_2^4 s_3^2}\geq 1$ (since $|a_{11}|$ must be at least 1). In that case, $\mathcal{B}$ is a union of two boxes in $\mathbb{R}^{10}$ each of whose sidelengths is bounded from below; by Lemma 10, we have

$$\sigma^*(S)\ll \mathrm{Vol}(\mathcal{B})O(X^{\epsilon})^2\ll \lambda^{-10}s_2^2s_3^4O(X^{\epsilon})$$

so that

$$N^*(S;X)\ll O(X^{\epsilon})\int_{\lambda=X^{-\frac{1}{12}}}^{c}\int_{s_1,s_2,s_3=\frac12}^{\infty}\lambda^{-10}s_1^{-2}s_2^{-4}s_3^{-2}\,d^\times s\,d^\times\lambda=O(X^{10/12+\epsilon}).$$

We now consider the cases where exactly two of r,s,t are equal to zero. This condition implies that either $a_{11}=b_{11}=0$ (which does not occur by hypothesis), $a_{22}=b_{22}=0$, or $a_{33}=b_{33}=0$.

If $a_{33}=b_{33}=0$, then the estimate

$$\prod_{t\in T\setminus\{a_{33},b_{33}\}} t=O(X^{10/12}) \tag{26}$$

(again with variables equal to zero replaced by a_{11}) shows that there are at most $O(X^{10/12+\epsilon})$ possibilities for the variables in T.

Finally, suppose $a_{22}=b_{22}=0$. We show that as v ranges over H, on average one expects $O(X^{10/12})$ values for (A,B) in this case. Let S denote the set of $(A,B)\in V_{\mathbb{Z}}$ for which $a_{22}=b_{22}=0$. Then we have as before the estimate (11) for $N^*(S;X)$. The value of $\sigma^*(S)$ is the number of integer points in the region defined by (10) together with the condition $a_{22}=b_{22}=0$. As before, this defines a region $\mathcal{B}$ in $\mathbb{R}^{10}$ which—whenever it has an integer point—becomes the union of two boxes whose edges are parallel to the coordinate axes and whose lengths are bounded from below. Now $\mathrm{Vol}(\mathcal{B})\ll\lambda^{-10}s_2^{-4}s_3^4$, so by Lemma 10, we obtain

$$N^*(S;X)\ll\int_{\lambda=X^{-\frac{1}{12}}}^{c}\int_{s_1,s_2,s_3=\frac12}^{\infty}\lambda^{-10}s_1^{-2}s_2^{-10}s_3^{-2}\,d^\times s\,d^\times\lambda=O(X^{10/12}).$$

This completes the proof of Lemma 13. □

2.5. *Cutting the cusps.* Let $0<\delta<\frac{1}{12}$.

LEMMA 14. *Let v take a random value in H uniformly with respect to the measure $|\mathrm{Disc}(v)|^{-1}\,dv$. Then the expected number of $(A,B)\in\mathcal{F}v$ with $|\mathrm{Disc}(A,B)|<X$ such that $0<|a_{11}|<X^{\delta}$ is $O(X^{11/12+\delta})$.*

Proof. We partition $V_{\mathbb{R}}$ into $\cup V(m)$, where $V(m)$ denotes the subset of $V_{\mathbb{R}}$ such that $|a_{11}| = m$. To handle $N^*(V(m);X)$ for $m \geq 1$, we use again the estimate (11). In this case, the quantity $\sigma^*(V(m))$ is equal to the number of integer points (A,B) satisfying the inequalities (10) and the condition that $|a_{11}| = m$. This set of integer points can be nonempty only if $\frac{60}{s_1 s_2^4 s_3^2}$ is at least m. In that case, the region $\mathcal{B}$ defined by (10) and $|a_{11}| = m$ is the union of two 11-dimensional boxes (contained in the hyperplanes of $V_{\mathbb{R}}$ defined by $a_{11} = \pm m$) whose sidelengths are all bounded below by an absolute constant. By Lemma 10,

$$\sigma^*(V(m)) \ll \mathrm{Vol}(\mathcal{B}) \ll \lambda^{-11} s_1 s_2^4 s_3^2.$$

Estimate (11) thus gives

$$N(V(m);X) \ll \int_{\lambda=X^{-\frac{1}{12}}}^{c} \int_{s_1,s_2,s_3=\frac{1}{2}}^{\infty} \lambda^{-11} s_1^{-1} s_2^{-2} s_3^{-4} d^\times s\, d^\times\lambda = O(X^{11/12})$$

where the implied constant is independent of m. Hence

$$N^*(\cup_{1\leq m\leq X^\delta} V(m);X) = \sum_{m=1}^{\lfloor X^\delta \rfloor} N(V(m);X) = X^\delta O(X^{11/12}) = O(X^{11/12+\delta}),$$

as desired. $\square$

Lemma 15. *Let v take any value in $H \cap V^{(i)}$. Let $\mathcal{R}_X = \mathcal{R}_X(v)$ denote the submultiset of points in $\mathcal{F}v$ having absolute discriminant less than X, and let $\mathcal{R}_X^{(\delta)} = \{(A,B) \in \mathcal{R}_X : |a_{11}| \geq X^\delta\}$. Then the number of integral elements in $\mathcal{R}_X^{(\delta)}$ is*

$$\mathrm{Vol}(\mathcal{R}_X^{(\delta)}) + O(X^{1-\delta+\epsilon}),$$

where $\mathrm{Vol}(\mathcal{R}_X^{(\delta)})$ *denotes the volume of the multiset $\mathcal{R}_X^{(\delta)}$.*

Proof. Let $\mathcal{R}_X^{(\delta)}$ be as in the statement of the lemma. Then it is easy to see that the region $\mathcal{R}_X^{(\delta)}$ is bounded; indeed, the conditions $|a_{11}| \geq X^\delta$ and $a_{11}^3 t = O(X^{1/3})$ imply that $t = O(X^{1/3-3\delta})$ for all $t \in T$. Furthermore, the various boundaries of $\mathcal{R}_X^{(\delta)}$ are defined by a bounded number of algebraic surfaces of bounded degree. By Lemma 9, it follows that the number of integer points in the multiset $\mathcal{R}_X^{(\delta)}$ is

$$\mathrm{Vol}(\mathcal{R}_X^{(\delta)}) + O(\mathrm{Vol}(\bar{\mathcal{R}}_X^{(\delta)})) \tag{27}$$

where $\mathrm{Vol}(\bar{\mathcal{R}}_X^{(\delta)})$ denotes the greatest r-dimensional volume of a projection of $\mathcal{R}_X^{(\delta)}$ onto any of the r-dimensional coordinate subspaces ($1 \leq r \leq 11$) in $V_{\mathbb{R}}$.

Let T again denote the set of twelve variables $\{a_{ij}, b_{ij}\}$, let T' be any proper subset of T, and consider the projection of $\mathcal{R}_X^{(\delta)}$ onto the coordinate hyperplane $Z_{T'}$ given by

$$Z_{T'} = \{t = 0 : t \in T \setminus T'\}.$$

We know by Lemma 8 that for $(A, B) \in \mathcal{R}_X^{(\delta)}$,

$$|a_{11}|^{12-|T'|} \cdot \Big|\prod_{t \in T'} t\Big| < C_1 X$$

for some constant C_1. Since $|a_{11}| \geq X^\delta$, and $12 - |T'| \geq 1$, it follows that

$$\Big|\prod_{t \in T'} t\Big| < C_1 X^{1-\delta}. \tag{28}$$

Furthermore, we have seen that $|a_{11}| \geq X^\delta$ implies that for any $t \in T'$,

$$|t| < C_2 X^{1/3} \tag{29}$$

for some constant C_2. Thus the projection of $\mathcal{R}_X^{(\delta)}$ onto $Z_{T'}$ is contained in the $|T'|$-dimensional region defined by (28) and (29). This region is seen to have volume at most

$$O(X^{1-\delta+\epsilon}),$$

for any proper subset $T' \subset T$.

Therefore, (27) implies that the number of integer points in $\mathcal{R}_X^{(\delta)}$ is given by

$$\mathrm{Vol}(\mathcal{R}_X^{(\delta)}) + O(X^{1-\delta+\epsilon}), \tag{30}$$

where the implied constant may be chosen independently of $v \in H \cap V^{(i)}$. This is the desired conclusion. □

LEMMA 16. *Let v take a random value in $H \cap V^{(i)}$ uniformly with respect to the measure $|\mathrm{Disc}(v)|^{-1}\,dv$, and let $\mathcal{R}_X = \mathcal{R}_X(v)$ and $\mathcal{R}_X^{(\delta)} = \mathcal{R}_X^{(\delta)}(v)$ be as in Lemma* 15. *Then the expected size of* $\mathrm{Vol}(\mathcal{R}_X) - \mathrm{Vol}(\mathcal{R}_X^{(\delta)})$ *is* $O(X^{11/12+\delta})$.

Proof. Let $E_i(X)$ denote the expected value of $\mathrm{Vol}(\mathcal{R}_X(v)) - \mathrm{Vol}(\mathcal{R}_X^{(\delta)}(v))$, as v varies over $H \cap V^{(i)}$. We may write

$$E_i(X) = \frac{1}{M_i} \int_{v \in V^{(i)}} \int_{\substack{x=(A,B)\in\mathcal{R}_X(v) \\ |a_{11}|<X^\delta}} \Phi(v)\,dx\,|\mathrm{Disc}(v)|^{-1} dv, \tag{31}$$

where both dv and dx denote Euclidean measure on $\mathbb{R}^{12}$. Let us denote by $V^{(i)}(\delta, X) \subset V^{(i)}$ the set $\{(A, B) \in V^{(i)} : |a_{11}| < X^\delta,\ |\mathrm{Disc}(A, B)| < X\}$.

Following (6)–(8) and the proof of Lemma 14, we then have

$$
\begin{aligned}
E_i(X) &= \frac{c'}{M_i}\int_{g\in\mathcal{F}^{-1}}\int_{x\in V^{(i)}(\delta,X)}\Phi(gx)\,dx\,dg \\
&\ll \int_{g\in KN'A'^{-1}\Lambda}\int_{x\in V^{(i)}(\delta,X)}\Phi(kna^{-1}\lambda x)\;\; s_1^{-2}s_2^{-6}s_3^{-6}\;\; dx\,d^\times\lambda\,d^\times s\,dn\,dk \\
&\ll \int_{\lambda,\,s_1,s_2,s_3}\int_{x\in V^{(i)}(\delta,X)}\Psi\big(\lambda\cdot a(s_1,s_2,s_3)^{-1}(A,B)\big)\;s_1^{-2}s_2^{-6}s_3^{-6}\,dx\,d^\times s\,d^\times\lambda \\
&\ll \int_{\lambda=X^{-\frac{1}{12}}}^{c}\int_{s_1,s_2,s_3=\frac{1}{2}}^{\infty}\int_{a_{11}=-X^\delta}^{X^\delta}\lambda^{-11}s_1^{-1}s_2^{-2}s_3^{-4}\;\;da_{11}\,d^\times s\,d^\times\lambda \\
&= O(X^{11/12+\delta}). \qquad\square
\end{aligned}
$$

Choose $\delta = 1/24$. Then Lemmas 11–16 yield

Proposition 17. *Let v take a random value in $H\cap V^{(i)}$ uniformly with respect to the measure $|\mathrm{Disc}(v)|^{-1}\,dv$, and let $\mathcal{R}_X=\mathcal{R}_X(v)$ denote the sub-multiset of points in $\mathcal{F}v$ having absolute discriminant less than X. Then the expected number of absolutely irreducible integral elements in $\mathcal{R}_X$ is*

$$\mathrm{Vol}(\mathcal{R}_X)+O(X^{23/24+\epsilon}).$$

Therefore, even though the total number of lattice points in $\mathcal{R}_X$ far exceeds the volume of $\mathcal{R}_X$ in general, the above proposition states that the number of *absolutely irreducible* lattice points in $\mathcal{R}_X$ will essentially be equal to the volume as $X\to\infty$.

2.6. *Computation of the fundamental volume.* To prove Theorem 7, it remains only to compute $\mathrm{Vol}(\mathcal{R}_X(v))$, where $\mathcal{R}_X(v)$ is defined as in Lemma 15. We will see that this volume depends only on whether v lies in $V^{(0)}$, $V^{(1)}$, or $V^{(2)}$; here $V^{(i)}$ again denotes the $G_{\mathbb{R}}$-orbit in $V_{\mathbb{R}}$ consisting of those elements (A,B) for which A and B possess $4-2i$ common zeros in $\mathbb{P}^2(\mathbb{R})$.

Before performing this computation, we state first some propositions regarding the group $G=\mathrm{GL}_2\times\mathrm{SL}_3$ and its 12-dimensional representation V.

Proposition 18. *The group $G_{\mathbb{R}}$ acts transitively on $V^{(i)}$, and the isotropy groups for $v\in V^{(i)}$ are given as follows*:

(i) S_4, *if* $v\in V^{(0)}$;

(ii) $C_2\times C_2$, *if* $v\in V^{(1)}$; *and*

(iii) D_4, *if* $v\in V^{(2)}$.

In view of Proposition 18, it is convenient to use the notation n_i to denote the order of the stabilizer of any vector $v \in V^{(i)}$. Proposition 18 implies that we have $n_0 = 24$, $n_1 = 4$, and $n_2 = 8$.

Now define the usual subgroups K, A_+, N, and $\bar{N}$ of $G_{\mathbb{R}}$ as follows:

$$\begin{aligned}
K &= \{\text{orthogonal transformations in } G_{\mathbb{R}}\};\\
A_+ &= \{a(t) : t \in \mathbb{R}_+^{\times 4}\}, \text{ where } a(t) = \left(\begin{pmatrix} t_1 & \\ & t_2 \end{pmatrix}, \begin{pmatrix} t_3 & & \\ & t_4 & \\ & & (t_3t_4)^{-1} \end{pmatrix} \right);\\
N &= \{n(u) : u \in \mathbb{R}^4\}, \text{ where } n(u) = \left(\begin{pmatrix} 1 & \\ u_1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & & \\ u_2 & 1 & \\ u_3 & u_4 & 1 \end{pmatrix} \right);\\
\bar{N} &= \{\bar{n}(x) : x \in \mathbb{R}^4\}, \text{ where } \bar{n}(x) = \left(\begin{pmatrix} 1 & x_1 \\ & 1 \end{pmatrix}, \begin{pmatrix} 1 & x_2 & x_3 \\ & 1 & x_4 \\ & & 1 \end{pmatrix} \right).
\end{aligned}$$

It is well-known that the natural product map $K \times A_+ \times N \to G_{\mathbb{R}}$ is an analytic diffeomorphism. In fact, for any $g \in G_{\mathbb{R}}$, there exist unique $k \in K$, $a = a(t_1, \ldots, t_4) \in A_+$, and $n = n(u_1, \ldots, u_4) \in N$ such that $g = k\,a\,n$. In particular, the element $\bar{n}(x) \in \bar{N}$ can also be factored uniquely in this way; the corresponding value of a is provided in the following proposition.

Proposition 19. *Let $\bar{n}(x_1, \ldots, x_4) \in \bar{N}$. Set*

$$q = 1 + x_1^2, \quad r = 1 + x_2^2 + (x_2x_4 - x_3)^2, \quad s = 1 + x_3^2 + x_4^2.$$

Then $\bar{n} = k\,a(t_1, t_2, t_3, t_4)\,n$, where

$$t_1 = 1/\sqrt{q}, \quad t_2 = \sqrt{q}, \quad t_3 = 1/\sqrt{r}, \quad t_4 = \sqrt{r}/\sqrt{s}.$$

Define an invariant measure dg on $G_{\mathbb{R}}$ as follows. Choose an invariant measure dk on K so that $\int_K 1\,dk = 1$, and define

$$\begin{aligned}
\int_{G_{\mathbb{R}}} f(g)dg &= \int_K \int_{\mathbb{R}^4} \int_{\mathbb{R}_+^{\times 4}} f(kna)d^\times t\,du\,dk\\
&= \int_K \int_{\mathbb{R}^4} \int_{\mathbb{R}_+^{\times 4}} t_1^{-1}t_2t_3^{-4}t_4^{-2} f(kan)d^\times t\,du\,dk.
\end{aligned}$$

Let $dy = dy_1\,dy_2 \cdots dy_{12}$ be the standard Euclidean measure on $V_{\mathbb{R}}$.

Proposition 20. *For any $f \in L^1(G_{\mathbb{R}})$,*

$$\int_{G_{\mathbb{R}}} f(g)dg = \frac{1}{32\pi^3} \int_{\mathbb{R}^{\times 4}} \int_{\mathbb{R}^4} \int_{\mathbb{R}^4} f(\bar{n}(x)n(u)a(t))dx\,du\,d^\times t.$$

Proof. We apply Proposition 19 to change variables, using the value of the definite integral

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{qrs}\, dx_1\, dx_2\, dx_3\, dx_4 \;=\; 2\pi^3. \qquad \square$$

PROPOSITION 21. *For $i = 0$, 1, or 2, let $f \in C_0(V^{(i)})$, and let y denote any element of $V^{(i)}$. Then*

$$\int_{g\in G_{\mathbb{R}}} f(g\cdot y)dg = \frac{n_i}{6\pi^3}\int_{v\in V^{(i)}} |\mathrm{Disc}(v)|^{-1} f(v)\, dv.$$

Proof. It suffices to prove the equality for

$$y = \left(\begin{bmatrix} 1 & & \\ & -1 & \\ & & \end{bmatrix}, \begin{bmatrix} & & \\ & -1 & \\ & & 1 \end{bmatrix}\right) \in V^{(0)},$$

$$y = \left(\begin{bmatrix} 1 & & \\ & 1 & \\ & & -1 \end{bmatrix}, \begin{bmatrix} & & \\ & & 1 \\ & 1 & \end{bmatrix}\right) \in V^{(1)}, \text{ or}$$

$$y = \left(\begin{bmatrix} 1 & & \\ & 1 & \\ & & \end{bmatrix}, \begin{bmatrix} & & \\ & 1 & \\ & & 1 \end{bmatrix}\right) \in V^{(2)}.$$

Put

$$(z_1, \ldots, z_{12}) = \bar{n}(x)n(u)a(t)\cdot y.$$

Then the form $\mathrm{Disc}(z)^{-1}dz_1 \wedge \cdots \wedge dz_{12}$ is a $G_{\mathbb{R}}$-invariant measure, and so we must have

$$\mathrm{Disc}(z)^{-1}dz_1 \wedge \cdots \wedge dz_{12} = c\; dx \wedge du \wedge d^{\times}t$$

for some constant factor c. An explicit calculation shows that $c = -3/16$ in all three cases. By Proposition 18, $G_{\mathbb{R}}$ is an n_i-fold covering of $V^{(i)}$ via the map $g \to g\cdot y$, where $n_i = 24$, 4, or 8 for $i = 0$, 1, or 2 respectively. Hence

$$\begin{aligned}\int_{G_{\mathbb{R}}} f(g\cdot y)dg &= n_i \cdot \frac{1}{32\pi^3}\cdot\frac{16}{3}\int_{V^{(i)}} |\mathrm{Disc}(v)|^{-1} f(v)dv \\ &= \frac{n_i}{6\pi^3}\int_{V^{(i)}} |\mathrm{Disc}(v)|^{-1} f(v)dv,\end{aligned}$$

as desired. $\square$

Finally, for a vector $v_i \in V^{(i)}$, we obtain using Proposition 21 that

$$\begin{aligned}\mathrm{Vol}(\mathcal{R}_X(v_i)) &= \frac{6\pi^3}{n_i}\int_0^{X^{1/6}} t^6 d^\times t \;\cdot \int_{G_{\mathbb{Z}}\backslash G^1_{\mathbb{R}}} dg \\ &= \frac{6\pi^3}{n_i}\cdot\frac{X}{6}\cdot\frac{\zeta(2)}{\pi}\cdot\frac{\zeta(2)\zeta(3)}{2\pi^2} \;=\; \frac{\zeta(2)^2\zeta(3)}{2n_i}X,\end{aligned}$$

proving Theorem 7.

3. Pairs of ternary quadratic forms and Theorems 1–5

Theorem 6 and Theorem 7 together now immediately imply the following

THEOREM 22. *Let $M_4^{*(i)}(\xi,\eta)$ denote the number of isomorphism classes of pairs (Q,R) such that Q is an order in an S_4-quartic field with $4-2i$ real embeddings, R is a cubic resolvent ring of Q, and $\xi < \mathrm{Disc}(Q) < \eta$. Then*

$$\begin{aligned}&\text{(a)}\quad \lim_{X\to\infty}\frac{M_4^{*(0)}(0,X)}{X} &&= \frac{\zeta(2)^2\zeta(3)}{48};\\ &\text{(b)}\quad \lim_{X\to\infty}\frac{M_4^{*(1)}(-X,0)}{X} &&= \frac{\zeta(2)^2\zeta(3)}{8};\\ &\text{(c)}\quad \lim_{X\to\infty}\frac{M_4^{*(2)}(0,X)}{X} &&= \frac{\zeta(2)^2\zeta(3)}{16}.\end{aligned}$$

To obtain finer asymptotic information on the distribution of quartic rings (in particular, without the weighting by the number of cubic resolvents), we need to be able to count absolutely irreducible $G_{\mathbb{Z}}$-equivalence classes in $V_{\mathbb{Z}}$ lying in certain subsets $S \subset V_{\mathbb{Z}}$. If S is defined, say, by *finitely many* congruence conditions, then this can easily be done; we have

$$\lim_{X\to\infty}\frac{N(S\cap V^{(i)};X)}{X} = \frac{\zeta(2)^2\zeta(3)}{2n_i}\prod_p \mu_p(S), \tag{32}$$

where $\mu_p(S)$ denotes the p-adic density of S in $V_{\mathbb{Z}}$, and $n_i = 24$, 4, or 8 for $i = 0$, 1, or 2 respectively. This refinement of Theorem 7 is proved in exactly the same way as the original theorem.

We recall from [3, §4.10], however, that the set S of elements (A,B) in $V_{\mathbb{Z}}$ corresponding to maximal orders is defined by infinitely many congruence conditions. To prove that (32) still holds for such a set, we require a uniform estimate on the error term when only finitely many factors are taken in (32). This estimate is provided in Section 3.2. In Section 3.3, we then use the estimate to complete the proofs of Theorems 1–5.

3.1. *Nowhere overramified quartic fields.* Let Q be an order in an S_4-quartic field, and let $p \in \mathbb{Z}$ be a prime such that Q is maximal at p. We say p is *overramified* in Q if (p) factors into primes in Q as P^4, P^2, or $P_1^2P_2^2$; similarly, the archimedean prime of $\mathbb{Z}$ (the "prime at infinity") is *overramified* in Q if it factors into the product of two ramified places (i.e., if Q is totally complex). A maximal quartic order Q (or the quartic field K_4 in which it lies) is *nowhere overramified* if no prime of $\mathbb{Z}$ (finite or infinite) is overramified in Q.

The significance of being "nowhere overramified" is as follows. Given an S_4-quartic field K_4, let K_{24} denote its Galois closure. Let K_3 denote a cubic field contained in K_{24} (the "cubic resolvent field"), and let K_6 be the unique quadratic extension of K_3 such that the Galois closure of K_6 over $\mathbb{Q}$ is K_{24}. Then one checks that the quadratic extension K_6/K_3 is unramified precisely when the quartic field K_4 is nowhere overramified. Conversely, if K_3 is a noncyclic cubic field, and K_6 is an unramified quadratic extension of K_3, then the Galois closure of K_6 is an S_4-extension K_{24} which contains up to conjugacy a unique, nowhere overramified quartic extension K_4.

3.2. *A uniformity estimate.* Let us denote by $\mathcal{V}_p$ the set of all $(A, B) \in V_{\mathbb{Z}}$ corresponding to quartic orders Q that are maximal at p and in which p is not overramified. Let $\mathcal{W}_p = V_{\mathbb{Z}} - \mathcal{V}_p$. In order to apply a simple sieve to obtain Theorems 1–5, we require the following proposition, analogous to Proposition 1 in [15] (though our proof is significantly simpler).

Proposition 23. $N(\mathcal{W}_p; X) = O(X/p^2)$, *where the implied constant is independent of* p.

Proof. The set $\mathcal{W}_p$ may be naturally partitioned into two subsets: $\mathcal{W}_p^{(1)}$, the set of points $(A, B) \in V_{\mathbb{Z}}$ corresponding to quartic rings not maximal at p; and $\mathcal{W}_p^{(2)}$, the set of points $(A, B) \in V_{\mathbb{Z}}$ corresponding to quartic rings that are maximal at p but also overramified at p.

We first treat $\mathcal{W}_p^{(1)}$. We will need the following lemma.

Lemma 24. *The number of maximal S_4-quartic orders* (*equivalently, the number of S_4-quartic fields*) *having absolute discriminant less than X is $O(X)$.*

Lemma 24 follows immediately from Theorem 22, since we have shown previously that every quartic ring has at least one cubic resolvent ring ([3, Corollary 4]).

To estimate $N(W_p^{(1)}; X)$ using Lemma 24, we wish to know for any multiple k of p that (a) the number of subrings of index k in a maximal quartic ring Q is not too large relative to k; and (b) the number of cubic resolvents that such a subring can possess is also not too large relative to k. For (a), a much stronger result than we need has been proved by Nakagawa [20], whose methods imply that the number of orders having index $k = \prod p_i^{e_i}$ in a maximal quartic ring Q is at most $O(\prod p_i^{(2+\epsilon)\lfloor e_i/4\rfloor})$, where $\prod p_i^{e_i}$ denotes the prime power decomposition of k. Any such order will of course have discriminant $k^2\mathrm{Disc}(Q)$.

As for (b), we have shown in [3, Corollary 4] that the number of cubic resolvents of a quartic ring having content n is $\sigma(n)$, where σ denotes the usual sum-of-divisors function. (Recall that the *content* of a quartic ring Q is the largest integer n such that $Q = \mathbb{Z} + nQ'$ for some quartic ring Q'.) In particular, rings having content 1 possess a single cubic resolvent.

Since every content n quartic ring Q arises as $\mathbb{Z}+nQ'$ for a unique content 1 quartic ring Q', and $\mathrm{Disc}(Q) = n^6\,\mathrm{Disc}(Q')$, we conclude

$$N(\mathcal{W}_p^{(1)};X) < \sum_{n=1}^{\infty}\frac{\sigma(n)}{n^6}\Big(\sum_{e=1}^{\infty}\frac{p^{(2+\epsilon)\lfloor e/4\rfloor}}{p^{2e}}\Big)\prod_{q\neq p}\Big(\sum_{e=0}^{\infty}\frac{q^{(2+\epsilon)\lfloor e/4\rfloor}}{q^{2e}}\Big)O(X) = O(X/p^2),$$

as desired.

We turn next to $\mathcal{W}_p^{(2)}$. Let us say a quadratic extension K_6 of a noncyclic cubic field K_3 is *acceptable* if the Galois closure of K_6 over $\mathbb{Q}$ has Galois group a transitive subgroup of S_4. For a fixed K_3, let $g(n)$ denote the number of acceptable quadratic extensions whose conductor has absolute norm n. To estimate $g(n)$, we require two lemmas. The first lemma is due to Baily [1]:

Lemma 25 (Baily). *K_6 is an acceptable quadratic extension of K_3 if and only if $N_{K_3/\mathbb{Q}}\mathrm{Disc}(K_6/K_3)$ is the square of an ideal in $\mathbb{Z}$.*

The next lemma gives an upper bound on the sum of $h_2^*(K_3)$ over all cubic fields K_3 having absolute discriminant less than X.

Lemma 26. *We have*

$$\sum_{K_3} h_2^*(K_3) \;=\; O(X), \tag{33}$$

where the sum ranges over all cubic fields K_3 having absolute discriminant less than X.

Lemma 26 follows from Lemma 24 as Theorem 5 will follow from Theorem 1.

Now it is a consequence of Lemma 25 that $g(n) = 0$ for nonsquare n. On the other hand, for a square integer $n = m^2$, class field theory implies that

$$g(m^2) \;<\; \kappa\, h_2^*(K_3)\, 3^{\omega(m)}, \tag{34}$$

where $\omega(m)$ denotes the number of prime factors of m, and κ is a constant bounded independently of K_3 (it corresponds to the even and infinite places; see [1] for details).

Lemma 26 and (34) now imply that, for some constant c'',

$$\begin{aligned} N(\mathcal{W}_p^{(2)};X) &\le \kappa \sum_{K_3} \sum_{\substack{p\mid m \\ m^2|\mathrm{Disc}(K_3)|<X}} 3^{\omega(m)} h_2^*(K_3) \\ &\le 3\kappa \sum_{p^2m^2<X} 3^{\omega(m)} \sum_{|\mathrm{Disc}(K_3)|<X/(p^2m^2)} h_2^*(K_3) \\ &\le 3\kappa c'' \frac{X}{p^2} \sum_{p^2m^2<X} \frac{3^{\omega(m)}}{m^2} \\ &< 3\kappa c'' \frac{X}{p^2} \sum_m \frac{3^{\omega(m)}}{m^2}. \end{aligned}$$

As the last sum converges absolutely, this concludes the proof of the proposition. □

3.3. *Proofs of Theorems* 1–5.

Proof of Theorem 1. As in [3], let $\mathcal{U}_p$ denote the set of all $(A,B)\in V_{\mathbb{Z}}$ that correspond to pairs (Q,R) for which Q is maximal at p, and let $\mathcal{U}=\cap_p\mathcal{U}_p$. Then $\mathcal{U}$ is the set of $(A,B)\in V_{\mathbb{Z}}$ corresponding to maximal quartic rings Q. In [3, Lemma 23], we determined the p-adic density $\mu_p(\mathcal{U}_p)$ of $\mathcal{U}_p$:

$$\mu_p(\mathcal{U}_p) = (p-1)^4\, p\, (p+1)^2\, (p^2+p+1)(p^3+p^2+2p+1)/\,p^{12}. \tag{35}$$

Suppose Y is any positive integer. It follows from (32) and (35) that

$$\lim_{X\to\infty} \frac{N^{(i)}(\cap_{p<Y}\mathcal{U}_p;X)}{X} = \frac{\zeta(2)^2\zeta(3)}{2n_i} \prod_{p<Y} [p^{-12}p\,(p^2-1)^2(p^3-1)(p^4+p^2-p-1)],$$

where we use the notation $N^{(i)}(S;X)$ for $N(S\cap V^{(i)};X)$. Letting Y tend to infinity, we obtain immediately that

$$\begin{aligned} \limsup_{X\to\infty} \frac{N^{(i)}(\mathcal{U};X)}{X} &\le \frac{\zeta(2)^2\zeta(3)}{2n_i} \prod_p [p^{-12}p\,(p^2-1)^2(p^3-1)(p^4+p^2-p-1)] \\ &= \frac{\zeta(2)^2\zeta(3)}{2n_i} \prod_p [(1-p^{-2})^2(1-p^{-3})(1+p^{-2}-p^{-3}-p^{-4})] \\ &= \frac{1}{2n_i} \prod_p (1+p^{-2}-p^{-3}-p^{-4}). \end{aligned}$$

To obtain a lower bound for $N^{(i)}(\mathcal{U};X)$, we note that

$$\bigcap_{p<Y} \mathcal{U}_p \subset (\mathcal{U} \cup \bigcup_{p\ge Y} \mathcal{W}_p).$$

Hence by Proposition 23,

$$\lim_{X\to\infty} \frac{N^{(i)}(\mathcal{U};X)}{X} \geq \frac{\zeta(2)^2\zeta(3)}{2n_i} \prod_{p<Y} [p^{-12} p\,(p^2-1)^2(p^3-1)(p^4+p^2-p-1)] - O(\sum_{p\geq Y} p^{-2}).$$

Letting Y tend to infinity completes the proof. □

Proof of Theorem 2. We first prove the analogue of Theorem 2 for the set $\mathcal{S}$ of S_4-quartic orders having content 1 (i.e., the *primitive* quartic orders); on such quartic rings the correspondence of Theorem 6 is bijective. To this end, let $\mathcal{S}_p$ denote the set of elements $(A,B)\in V_{\mathbb{Z}}$ having content prime to p, so that $\mathcal{S}=\cap_p\mathcal{S}_p$. Then as noted in [3], an element $(A,B)\in V_{\mathbb{F}_p}$ corresponds to a quartic ring with content prime to p if and only if A and B are linearly independent over $\mathbb{F}_p$. It follows that

$$\mu_p(\mathcal{S}_p) = (p^6-1)(p^6-p)/p^{12}.$$

The same argument as in the proof of Theorem 1 then shows that

$$\lim_{X\to\infty} \frac{N^{(i)}(\mathcal{S};X)}{X} = \frac{\zeta(2)^2\zeta(3)}{2n_i}\prod_p(1-p^{-5})(1-p^{-6}) = \frac{\zeta(2)^2\zeta(3)}{2n_i\,\zeta(5)\zeta(6)}. \tag{36}$$

To obtain Theorem 2 from (36), we observe that every content 1 ring Q_1 contains the content n ring $Q_n=\mathbb{Z}+nQ_1$, and conversely, every content n ring Q_n arises from a unique content 1 ring Q_1 in this way. Furthermore, if Q_1 has discriminant D then Q_n has discriminant n^6D. It follows that

$$\lim_{X\to\infty}\frac{M_4^{(i)}(0,X)}{X} = \sum_{n=1}^{\infty}\frac{1}{n^6}\frac{\zeta(2)^2\zeta(3)}{2n_i\zeta(5)\zeta(6)} = \frac{\zeta(2)^2\zeta(3)}{2n_i\,\zeta(5)},$$

as desired. □

Proof of Theorem 3. It is known that the Artin symbol (K_{24}/p) equals $\langle e\rangle$, $\langle(12)\rangle$, $\langle(123)\rangle$, $\langle(1234)\rangle$, and $\langle(12)(34)\rangle$ precisely when the splitting type of p in Q is (1111), (112), (13), (4), or (22) respectively, where Q denotes the ring of integers in K_4. As in [3], let $T_p(\sigma)$ denote the set of all $(A,B)\in V_{\mathbb{Z}}$ that correspond to quartic rings Q having a specified splitting type σ at p; then the set of all such $(A,B)\in V_{\mathbb{Z}}$ corresponding to maximal quartic rings Q is given by $\mathcal{U}\cap T_p(\sigma)$. Hence by the same argument as in the proof of Theorem 1, we have

$$\lim_{X\to\infty} X^{-1}N^{(i)}(\mathcal{U}\cap T_p(\sigma);X) = \mu_p(T_p(\sigma))\prod_{q\neq p}\mu_q(\mathcal{U}_q).$$

On the other hand, Lemma 21 of [3] gives the p-adic densities of $T_p(\sigma)$ for all splitting and ramification types σ. In particular, the values of $\mu_p(T_p(\sigma))$ for $\sigma=$ (1111), (112), (13), (4), or (22) occur in the ratio 1:6:8:6:3 for any value of p; this is the desired result. □

Proof of Theorem 4. By Theorem 1, the number of S_4-quartic fields having $4-2i$ real embeddings and absolute discriminant at most X is asymptotic to $c_i X$, where

$$c_0 = .0253477143\ldots, \quad c_1 = .1520862858\ldots, \quad c_2 = .0760431429\ldots.$$

Thus the total number of S_4-quartic fields with absolute discriminant at most X is asymptotic to cX, where $c = c_0 + c_1 + c_2 = .2534771431\ldots$.

The analogous constants for D_4-quartic fields, as obtained in [7], are given by

$$d_0 = .0049278439\ldots, \quad d_1 = .0098556878\ldots, \quad d_2 = .0375424793\ldots,$$

for a total of $d = .0523260112\ldots$.

It follows that the asymptotic proportion of S_4-quartic fields among all quartic fields having 4, 2, or 0 real embeddings, when ordering quartic fields by absolute discriminant, is given by $\frac{c_0}{c_0+d_0} \approx 83.723\%$, $\frac{c_1}{c_1+d_1} \approx 93.914\%$, and $\frac{c_2}{c_2+d_2} \approx 66.948\%$ respectively, for an overall proportion of $\frac{c}{c+d} \approx 82.889\%$. This yields Theorem 4. $\square$

Proof of Theorem 5. Let $\mathcal{V} = \cap_p \mathcal{V}_p$ be the set of all $(A,B) \in V_{\mathbb{Z}}$ corresponding to nowhere overramified maximal quartic rings. Using [3, Lemma 23], and the fact that $\mathcal{V}_p$ is simply the union of all $U_p(\sigma)$'s where $\sigma \neq (1^4)$, (2^2), or $(1^2 1^2)$, we obtain

$$\mu_p(\mathcal{V}_p) = p^{-12}\, p^2\, (p^2-1)^2\, (p^3-1)^2. \tag{37}$$

By the same argument as in Theorem 1, we therefore get

Lemma 27. *Let $L_4^{(i)}(\xi,\eta)$ denote the number of nowhere overramified S_4-quartic fields K having $4-2i$ real embeddings such that $\xi < \mathrm{Disc}(K) < \eta$. Then*

$$\text{(a)} \quad \lim_{X\to\infty} \frac{L_4^{(0)}(0,X)}{X} = \frac{\zeta(2)^2\zeta(3)}{48} \prod_p \mu_p(\mathcal{V}_p) = \frac{\zeta(2)^2\zeta(3)}{48}\, \zeta(2)^{-2}\zeta(3)^{-2} = 1/(48\zeta(3));$$

$$\text{(b)} \quad \lim_{X\to\infty} \frac{L_4^{(1)}(-X,0)}{X} = \frac{\zeta(2)^2\zeta(3)}{8} \prod_p \mu_p(\mathcal{V}_p) = \frac{\zeta(2)^2\zeta(3)}{8}\, \zeta(2)^{-2}\zeta(3)^{-2} = 1/(8\zeta(3)).$$

On the other hand, given a nowhere overramified S_4-quartic field K_4 with Galois closure K_{24}, we have observed earlier that in K_{24} is contained a unique (up to conjugacy) cubic field K_3 and a unique unramified extension K_6 of K_3.

In addition, the discriminant of K_4 is equal to the discriminant of K_3, and the number of quadruplets of quartic fields K_4 corresponding to a given K_3 in this way equals $h_2^*(K_3)-1$ (see Heilbronn [19] for full details). Therefore,

$$\begin{aligned} \sum_{0<\mathrm{Disc}(K_3)<X} (h_2^*(K_3)-1) &= L_4^{(0)}(0,X), \\ \sum_{-X<\mathrm{Disc}(K_3)<0} (h_2^*(K_3)-1) &= L_4^{(1)}(-X,0). \end{aligned} \tag{38}$$

Since Davenport and Heilbronn [15] have shown that

$$\begin{aligned} \lim_{X\to\infty} \frac{\sum_{0<\mathrm{Disc}(K_3)<X} 1}{X} &= 1/(12\zeta(3)), \\ \lim_{X\to\infty} \frac{\sum_{-X<\mathrm{Disc}(K_3)<0} 1}{X} &= 1/(4\zeta(3)), \end{aligned} \tag{39}$$

we conclude

$$\begin{aligned} \lim_{X\to\infty} \frac{\sum_{0<\mathrm{Disc}(K_3)<X} h_2^*(K_3)}{\sum_{0<\mathrm{Disc}(K_3)<X} 1} &= 1 + \lim_{X\to\infty} \frac{L_4^{(0)}(0,X)}{\sum_{0<\mathrm{Disc}(K_3)<X} 1} \\ &= 1 + \frac{1/(48\zeta(3))}{1/(12\zeta(3))} = \frac{5}{4}, \\ \lim_{X\to\infty} \frac{\sum_{-X<\mathrm{Disc}(K_3)<0} h_2^*(K_3)}{\sum_{-X<\mathrm{Disc}(K_3)<0} 1} &= 1 + \lim_{X\to\infty} \frac{L_4^{(1)}(-X,0)}{\sum_{-X<\mathrm{Disc}(K_3)<0} 1} \\ &= 1 + \frac{1/(8\zeta(3))}{1/(4\zeta(3))} = \frac{3}{2}. \end{aligned}$$

□

Acknowledgments. This article is based on Chapter 5 of the author's Ph.D. thesis [2] at Princeton University. I am extremely grateful to my advisor Andrew Wiles and to Peter Sarnak for all their enthusiasm, encouragement, and guidance during this work. I am also very thankful to Karim Belabas, Henri Cohen, Keith Conrad, Kiran Kedlaya, Hendrik Lenstra, Igor Rodnianski, Jean-Pierre Serre, and Don Zagier for many helpful conversations and comments on earlier versions of this manuscript.

I extend my gratitude to the Hertz Foundation for funding this work, and to the Clay Mathematics Institute for their subsequent support.

Clay Mathematics Institute, Cambridge, MA
Princeton University, Princeton, NJ
E-mail address: bhargava@math.princeton.edu

References

[1] A. M. Baily, On the density of discriminants of quartic fields, *J. Reine Angew. Math.* **315** (1980), 190–210.

[2] M. Bhargava, Higher Composition Laws, Ph.D. Thesis, Princeton University, June 2001.

[3] ———, Higher composition laws III: The parametrization of quartic rings, *Ann. of Math.* **159** (2004), 1329–1360.

[4] ———, Higher composition laws IV: The parametrization of quintic rings, *Ann. of Math.*, to appear.

[5] ———, The density of discriminants of quintic rings and fields, *Ann. of Math.*, to appear.

[6] H. Cohen, F. Diaz y Diaz, and M. Olivier, A survey of discriminant counting, in *Algorithmic Number Theory* (Sydney, 2002), *Lecture Notes in Comput. Sci*. **2369**, 80–94, Springer-Verlag, Berlin, 2002.

[7] H. Cohen, F. Diaz y Diaz, and M. Olivier, Enumerating quartic dihedral extensions of $\mathbb{Q}$, *Compositio Math.* **133** (2002), 65–93.

[8] H. Cohen and J. Martinet, Étude heuristique des groupes de classes des corps de nombres, *J. Reine Angew. Math.* **404** (1990), 39–76.

[9] ———, Heuristics on class groups: some good primes are not too good, *Math. Comp.* **63** (1994), 329–334.

[10] H. Cohn, The density of abelian cubic fields, *Proc. Amer. Math. Soc.* **5** (1954), 476–477.

[11] B. Datskovsky and D. J. Wright, The adelic zeta function associated to the space of binary cubic forms II: Local theory, *J. Reine Angew. Math.* **367** (1986), 27–75.

[12] H. Davenport, On a principle of Lipschitz, *J. London Math. Soc.* **26** (1951), 179–183.

[13] ———, On the class-number of binary cubic forms I and II, *J. London Math. Soc.* **26** (1951), 183–198.

[14] ———, Corrigendum: "On a principle of Lipschitz", *J. London Math. Soc.* **39** (1964), 580.

[15] H. Davenport and H. Heilbronn, On the density of discriminants of cubic fields II, *Proc. Roy. Soc. London Ser. A* **322** (1971), 405–420.

[16] B. N. Delone and D. K. Faddeev, *The Theory of Irrationalities of the Third Degree*, *Translations of Mathematical Monographs* **10**, 1964, A.M.S., Providence, RI.

[17] V. Ennola and R. Turunen, On totally real cubic fields, *Math. Comp.* **44** (1985), 495–518.

[18] G. W. Fung and H. C. Williams, On the computation of a table of complex cubic fields with discriminant $D > -10^6$, *Math. Comp.* **55** (1990), 313–325.

[19] H. Heilbronn, On the 2-classgroup of cubic fields, *Studies in Pure Math.* (1971), 117–119.

[20] J. Nakagawa, *Orders of a Quartic Field*, *Mem. Amer. Math. Soc.* **122** (1996), no. 583, A.M.S., Providence, RI.

[21] ———, On the relations among the class numbers of binary cubic forms, *Invent. Math.* **134** (1998), 101–138.

[22] M. Sato and T. Kimura, A classification of irreducible prehomogeneous vector spaces and their relative invariants, *Nagoya Math. J.* **65** (1977), 1–155.

[23] M. Sato and T. Shintani, On zeta functions associated with prehomogeneous vector spaces, *Ann. of Math.* **100** (1974), 131–170.

[24] A. Seidenberg, A new decision method for elementary algebra, *Ann. of Math.* **60** (1954), 365–374.

[25] A. Tarski, *A Decision Method for Elementary Algebra and Geometry*, 2nd ed., revised, Univ. of California Press, Berkeley and Los Angeles, Calif., 1951.

[26] S. Wong, Automorphic forms on GL(2) and the rank of class groups, *J. reine Angew. Math.* **515** (1999), 125–153.

[27] D. J. Wright and A. Yukie, Prehomogeneous vector spaces and field extensions, *Invent. Math.* **110** (1992), 283–314.

[28] A. Yukie, *Shintani Zeta Functions*, *London Mathematical Society Lecture Note Series* **183**, Cambridge Univ. Press, Cambridge, 1993.

[29] ———, Density theorems for prehomogeneous vector spaces, preprint.

(Received June 7, 2004)