Levels of distribution for sieve problems in prehomogeneous vector spaces

Takashi Taniguchi and Frank Thorne

July 6, 2017

Abstract

In our companion paper [28], we developed an efficient algebraic method for computing the Fourier transforms of certain functions defined on prehomogeneous vector spaces over finite fields, and we carried out these computations in a variety of cases.

Here we develop a method, based on Fourier analysis and algebraic geometry, which exploits these Fourier transform formulas to yield *level of distribution* results, in the sense of analytic number theory. Such results are of the shape typically required for a variety of sieve methods. As an example of such an application we prove that there are $\gg \frac{X}{\log X}$ quartic fields whose discriminant is squarefree, bounded above by X, and has at most eight prime factors.

In this paper we will develop a general technique sufficient to prove the following:

Theorem 1 *There is an absolute constant $C_3 > 0$ such that for each $X > 0$, there exist $\geq (C_3 + o_X(1))\frac{X}{\log X}$ cubic fields K whose discriminant is squarefree, bounded above by X, and has at most 3 prime factors.*

Theorem 2 *There is an absolute constant $C_4 > 0$ such that for each $X > 0$, there exist $\geq (C_4 + o_X(1))\frac{X}{\log X}$ quartic fields K whose discriminant is squarefree, bounded above by X, and has at most 8 prime factors.*

Theorem 1 improves on a result of Belabas and Fouvry [3] (which in turn improved upon Belabas [1]), where they obtained the same result with 7 in place of our 3, and our methods are in large part a further development of their ideas. In [3] they remarked that introducing a *weighted sieve* would lower this 7 to 4; the further improvement to 3 comes from an improvement in the corresponding *level of distribution*. The application to quartic fields is, to our knowledge, new.

Besides the weighted sieve, the main ingredients of our method are unusually strong estimates for the relevant exponential sums (which we obtained in [28]), and a suitably adapted version of the recently established Ekedahl-Bhargava geometric sieve [10].

The method is designed to yield strong *level of distribution estimates* as inputs to *sieve methods for prehomogeneous vector spaces*. We consider the following setup:

1. A sieve begins with a set of objects $\mathcal{A}$ to be sieved. Here this is defined by a representation V of an algebraic group G which we assume to be *prehomogeneous*: the action of $G(\mathbb{C})$ on $V(\mathbb{C})$ has a Zariski-open orbit, defined by the nonvanishing of a polynomial which we call the *discriminant*. We assume here that $\mathrm{Disc}(gx) = \mathrm{Disc}(x)$ identically for all $g \in G(\mathbb{Z})$, and take as our set $\mathcal{A} = \mathcal{A}(X)$ the set of $G(\mathbb{Z})$-orbits $x \in V(\mathbb{Z})$ with $0 < |\mathrm{Disc}(x)| < X$.

   Prehomogeneous vector spaces are the subject of a wealth of parametrization theorems (see, e.g., [29, 4, 5, 6, 8]) and corresponding theoerems concerning the arithmetic objects being parametrized. We refer also to [12] for a large number of interesting *coregular* (not prehomogneous) examples, for which the methods of this paper (and perhaps also [28]) are likely to apply.

2. For each prime p we define a notion of an object $x \in \mathcal{A}$ being 'bad at p'; a typical application of sieve methods is to estimate or bound the number of $x \in \mathcal{A}$ which are not bad at any prime $p < Y$, for some parameter Y. For Theorems 1 and 2 we will take 'bad at p' to mean 'has discriminant divisible by p'. Other defintions of 'bad' are also important, for example the Davenport-Heilbronn nonmaximality condition of [17, 2, 13, 27].

   We may work with any definition of 'bad' meeting the following technical condition: there is an integer $a \geq 1$ and a $G(\mathbb{Z}/p^a\mathbb{Z})$-invariant subset $S_p \subseteq V(\mathbb{Z}/p^a\mathbb{Z})$, such that $x \in V(\mathbb{Z})$ is bad if and only if its reduction $\pmod{p^a}$ is in S_p.

3. For each squarefree integer q, we require estimates for the number of $x \in \mathcal{A}(X)$ bad at each prime dividing q. These may be obtained via the geometry of numbers (see, among other references, [17, 1, 3, 9, 13, 19]) or using Shintani zeta functions [27], and we develop a third (simpler) method here.

   Here these estimates will be of the form $C\omega(q)X^{r/d} + E(X,q)$ for a fixed constant C and multiplicative function ω, and an error term $E(X,q)$ which we want to bound.

4. In cases where $\mathcal{A}$ depends on a parameter X, a *level of distribution* is any $\alpha > 0$ for which the sum $\sum_{q \leq X^\alpha} |E(X,q)|$ is adequately small. Typically, and here, a cumulative error $\ll_A X(\log X)^{-A}$ (for each $A > 0$) is more than sufficient – but see [13, 27, 26] for examples where a power savings is relevant.

We refer to books such as [16] and [21] for examples of sieve methods and applications; many of the methods can be used essentially as black boxes, for which the level of distribution is the most important input. *It is the goal of this paper to develop a method for proving levels of distribution for prehomogeneous vector spaces which are quantitatively as strong as possible.*

Typically, an important ingredient is finite Fourier analysis. Let $\Psi_p : V(\mathbb{Z}/p^a\mathbb{Z}) \to \mathbb{C}$ be the characteristic function of the subset S_p described above, and let $\widehat{\Psi_p}$ be its Fourier transform (defined by (5) below). We expect upper bounds on $E(X,p)$ to follow from statements to the effect that that the function Ψ_p is equidistributed, and upper bounds on the L_1-norm of $\widehat{\Psi_p}$ constitute a strong quantitatve statement of this equidistribution.

The basic heuristic of this paper is that L_1 *norm bounds for Fourier transforms over finite fields should lead to level of distribution statements for arithmetic objects.* For examples carried out via the geometry of numbers we refer to [3], [13, (80)-(83)], and [20, Proposition 9.2]; for examples using the Shintani zeta function method we refer to [27] (especially Section 3) or the forthcoming work [24].

In the present paper we develop a simpler version of this idea, not requiring any knowledge of the geometry of the 'cusps' or of the analytic behavior of the zeta function, in the context of a *lower bound sieve.* For any Schwartz function ϕ, the Poisson summation formula takes the form

$$\sum_{x \in V(\mathbb{Z})} \Psi_q(x)\phi(xX^{-1/d}) = X^{\dim(V)/d} \sum_{w \in V^*(\mathbb{Z})} \widehat{\Psi_q}(w)\widehat{\phi}\left(\frac{wX^{1/d}}{q}\right), \tag{1}$$

and for suitable ϕ the left side will be a smoothed undercount of the number of $G(\mathbb{Z})$-orbits $x \in V(\mathbb{Z})$ with $0 < |\operatorname{Disc}(x)| < X$, satisfying the 'local conditions' described by Ψ_q. The left side of (1) takes the role of (the counting function of) $\mathcal{A}(X)$, and the function Ψ_q may be used to detect those x which are 'bad at q'.

On the right side of (14), the $w = 0$ term the expected main term, and the rapid decay of $\widehat{\phi}$ will imply that the error term is effectively bounded by a sum of $|\widehat{\Psi_q}(w)|$ over a box of side length $\ll qX^{-1/d}$.

Therefore we are reduced to bounding sums of $|\widehat{\Psi_q}(w)|$ over boxes and over ranges of squarefree q. Obviously we require bounds on the individual $|\widehat{\Psi_q}(w)|$ to proceed. At this point we could incorporate the general bounds of Fouvry and Katz [20]; indeed, our method would quite directly exploit the geometric structure of their results. However, in the cases of interest much stronger bounds hold for the $|\widehat{\Psi_q}(w)|$ than are proved in [20]. We proved this in [28] as a special case of exact formulas for $\widehat{\Psi_q}(w)$, for any of five prehomogeneous vector spaces V, any squarefree integer q, and any $G(\mathbb{Z}/q\mathbb{Z})$-invariant function Ψ_q.

Our L_1-norm heuristic arises by replacing each $|\widehat{\Psi_q}(w)|$ by its average over $V^*(\mathbb{Z}/q\mathbb{Z})$. In practice there are limits to the equidistribution of $|\widehat{\Psi_q}(w)|$, and algebraic geometry now takes center stage. For $q = p \neq 2$ prime, the largest values of $|\widehat{\Psi_p}(w)|$ will be confined to $w \in \mathfrak{X}(\mathbb{F}_p)$ for a *scheme* $\mathfrak{X}$, defined over $\mathbb{Z}$, and of high codimension. (The same is also true of the Fouvry-Katz bounds.)

We now apply the *Ekedahl-Bhargava* geometric sieve [10], which essentially bounds the number of pairs (w, p) with w in our box and p prime. As we must work with general squarefree integers q, and a filtration of schemes $\mathfrak{X}_i$ rather than just one fixed $\mathfrak{X}$, we develop a variation of the geometric sieve adapted to this counting problem.

Organization of this paper. For simplicity we structure our paper around the proof of Theorem 2, even though our more important aim is to present a method which works in much greater generality. We begin in Section 1 by introducing the prehomogeneous vector space $(G, V) = (\mathrm{GL}_2 \times \mathrm{GL}_3, 2 \otimes \mathrm{Sym}_2(3))$ and describing its relevant properties. We take some care to delineate which of its properties are relevant to the proof, so that the reader can see what is required to adapt our method to other prehomogeneous vector spaces and to other sieve problems.

In Section 2 we introduce the *weighted sieve* of Richert [25] and Greaves [22], used to conclude Theorems 1 and 2, and we also precisely formulate the level of distribution statement which it will require.

In Section 3 we introduce our smoothing method. Our main result is Proposition 7, which states that a level of distribution follows from an essentially combinatorial estimate. The proof is a fairly typical application of Poisson summation, and follows along lines that should be familiar to experts. The proof is carried out in a general setting, and we state all of our assumptions at the beginning of the section.

Section 4 proves Theorem 1, and may be skipped without loss of continuity. Here the geometry is simple enough that we may conclude without introducing any algebro-geometric machinery, and the argument may be read as a prototype for the generalities which follow.

In Section 5 we introduce some algebraic geometry to describe the $G(\mathbb{F}_p)$-orbits on $V(\mathbb{F}_p)$ in terms of schemes defined over $\mathbb{Z}$. This sets up an application of the Ekedahl-Bhargava geometric sieve [10], of which we develop a variation in Section 6.

Section 7 is the heart of the proof, further developing the geometric sieve to prove our level of distribution statement. The proof is specalized to the particular (G, V) and Φ_q being studied, but the generalization to other cases should be immediate.

Finally, in Section 8 we prove Theorem 2. Essentially the proof is a formal consequence of our level of distribution, although there are a few technicalities pertaining to this particular (G, V) and its arithmetic interpretation.

Remark. The quartic fields produced by Theorem 2 will all have Galois group S_4, and the theorem still holds if we specify that $\mathrm{Disc}(K)$ should be positive or negative, or indeed that K has a fixed number of real embeddings.

There are other methods for producing almost-prime quartic field discriminants, One, suggested to us by Bhargava, is to specialize 11 of the 12 variables in V to particular values and then apply results on polynomials in one variable; a second is to apply results [15] on quartic fields with fixed cubic resolvent.

In either case we obtain quartic field discriminants with fewer than 8 prime factors. but it is unclear how to get, say, $\gg X^{9/10}$ of them, let alone $\gg X(\log X)^{-1}$. Moreover, proving that one obtains *fundamental* discriminants seems to be nontrivial with the first method, and the second method intrinsically produces non-fundamental discriminants.

Notation. We observe the following conventions in this paper. x will denote a general element of $V(\mathbb{Z})$ (sometimes only up to $G(\mathbb{Z})$-equivalence). r will denote the dimension of V and d will denote the degree of its (homogeneous) 'discriminant' polynomial. X will indicate a discriminant bound, and in Section 2 we write X and Y in place of the x and X of [21]. p will denote a prime and q a squarefree integer, in contrast to [28] where q was used for the cardinality of a finite field.

1 The 'quartic' representation and its essential properties

For each ring R (commutative, with unit), let $V(R) = R^2 \otimes \mathrm{Sym}_2(R^3)$ be the set of pairs of ternary quadratic forms with coefficients in R; when 2 is not a zero divisor in R, we also regard $V(R)$ as the set of pairs of 3×3 symmetric matrices. Let $G(R) := \mathrm{GL}_2(R) \times \mathrm{GL}_3(R)$; there is an action of G on V, defined by

$$(g_2, g_3) \cdot (A, B) = (r \cdot g_3 A g_3^T + s \cdot g_3 B g_3^T, t \cdot g_3 A g_3^T + u \cdot g_3 B g_3^T),$$

where $g_2 := \begin{pmatrix} r & s \\ t & u \end{pmatrix}$.

The *discriminant* is defined (see [6, p. 1340]) by the equation

$$\mathrm{Disc}((A, B)) := \mathrm{Disc}(4 \det(Ax + By)), \tag{2}$$

where $4 \det(Ax + By)$ is a binary cubic form in the variables x and y, and the second 'Disc' above is its discriminant.

It was proved by Bhargava [6] that the $G(\mathbb{Z})$-orbits on $V(\mathbb{Z})$ parametrize quartic rings, in a sense that we recall precisely in Section 8. (This parametrization, together with a geometry of numbers argument, allowed Bhargava to prove [7] an asymptotic formula for the number of quartic fields of bounded discriminant.)

We note the following additional facts about this representation. Although we will not attempt to axiomatize our method here, these are the inputs required to establish a lower bound sieve for $G(\mathbb{Z})$-orbits on $V(\mathbb{Z})$. (Arithmetic applications, such as passing from quartic rings to quartic *fields* as we do in Section 8, may in some cases require extra steps which will not generalize as readily.)

1. (*Homogeneity of the discriminant; definitions of r and d.*) By (2), Disc is a homogeneous polynomial of degree 12 on V. We write $d = 12$ for the degree of this polynomial, and $r := \dim(V) = 12$; these quantities coincide in this and other interesting cases, but not always.

2. (*Approximation of the fundamental domain.*) Let $\mathcal{F}$ be a fundamental domain for the action of $G(\mathbb{Z})$ on $V(\mathbb{R})$, and let $\mathcal{F}^1$ be the subset of $x \in \mathcal{F}$ with $0 < |\mathrm{Disc}(x)| < 1$. We approximate $\mathcal{F}^1$ by choosing a smooth (Schwartz class) function $\phi : V(\mathbb{R}) \to [0, 1]$ compactly supported within $\mathcal{F}^1$.

   Since the discriminant is homogeneous of degree d, the weighting function $\phi(xX^{-1/d})$ is a smoothed undercount of those $x \in G(\mathbb{Z})\backslash V(\mathbb{Z})$ with $0 < |\mathrm{Disc}(x)| < X$. That is, the role of the (counting function of the) set $\mathcal{A}(X)$ described in the introduction is taken by the expression

   $$\sum_{x \in V(\mathbb{Z})} \phi(xX^{-1/d}). \tag{3}$$

   Although it won't be necessary here, it is possible to approximate $G(\mathbb{Z})\backslash V(\mathbb{R})$ as closely as we wish in the sense that, for any $\beta < 1$, we may additionally require the locus $\mathcal{F}_\beta$ of x with $\phi(x) = 1$ to satisfy

   $$\frac{\mathrm{Vol}(\mathcal{F}_\beta)}{\mathrm{Vol}(\mathcal{F}^1)} > \beta. \tag{4}$$

   A number of variations are possible; for example we could restrict $\mathcal{F}^1$ to those $x \in \mathcal{F}$ with a particular sign, or choose ϕ to be supported away from any algebraic subset of $V(\mathbb{R})$ defined by the vanishing of one or more homogeneous equations.

3. (*Fourier transform formulas.*) For a squarefree integer q we let Ψ_q be the characteristic function of those $x \in V(\mathbb{Z})$ with $q \mid \mathrm{Disc}(x)$; this function factors through the reduction map $V(\mathbb{Z}) \to V(\mathbb{Z}/q\mathbb{Z})$. Its Fourier transform $\widehat{\Psi_q} \colon V^*(\mathbb{Z}/q\mathbb{Z}) \to \mathbb{C}$ is defined by the usual formula

   $$\widehat{\Psi_q}(x) = q^{-r} \sum_{x' \in V(\mathbb{Z}/q\mathbb{Z})} \Psi_q(x') \exp\left(\frac{2\pi i [x', x]}{q}\right). \tag{5}$$

The Fourier transform $\widehat{\Psi_q}$ is easily seen to be multiplicative in q, and for $q = p \neq 2$ we proved the following explicit formula in [28]:

(6)
$$\widehat{\Psi_p}(x) = \begin{cases} p^{-1} + 2p^{-2} - p^{-3} - 2p^{-4} - p^{-5} + 2p^{-6} + p^{-7} - p^{-8} & x \in \mathcal{O}_0, \\ p^{-3} - p^{-4} - 2p^{-5} + 2p^{-6} + p^{-7} - p^{-8} & x \in \mathcal{O}_{D1^2}, \\ 2p^{-4} - 5p^{-5} + 3p^{-6} + p^{-7} - p^{-8} & x \in \mathcal{O}_{D11}, \\ p^{-4} - 3p^{-5} + 2p^{-6} + p^{-7} - p^{-8} & x \in \mathcal{O}_{C\mathrm{s}}, \\ -p^{-5} + p^{-6} + p^{-7} - p^{-8} & x \in \mathcal{O}_{D2}, \mathcal{O}_{D\mathrm{ns}}, \mathcal{O}_{C\mathrm{ns}}, \mathcal{O}_{T11}, \mathcal{O}_{T2}, \\ -p^{-6} + 2p^{-7} - p^{-8} & x \in \mathcal{O}_{1^2 1^2}, \\ p^{-6} - p^{-8} & x \in \mathcal{O}_{2^2}, \\ p^{-7} - p^{-8} & x \in \mathcal{O}_{1^4}, \mathcal{O}_{1^3 1}, \mathcal{O}_{1^2 11}, \mathcal{O}_{1^2 2}, \\ -p^{-8} & x \in \mathcal{O}_{1111}, \mathcal{O}_{112}, \mathcal{O}_{22}, \mathcal{O}_{13}, \mathcal{O}_4. \end{cases}$$

When $p \neq 2$ there are 20 orbits for the action of $G(\mathbb{F}_p)$ on $V^*(\mathbb{F}_p)$, and each has a description that is essentially uniform in p; these are denoted by the $\mathcal{O}$ above, or by $\mathcal{O}(p)$ when we indicate the prime p explicitly. We refer to [28] for descriptions of each of the $\mathcal{O}$, together with computations of their cardinalities.

The L_1 norm of $\widehat{\Psi_p}(x)$ is $O(p^4)$ – better than square root cancellation! In particular the larger contributions come from the more singular orbits, and *our methods are designed to exploit this structure.*

What is required in general is that Ψ_p be any bounded function, which factors through the reduction map $V(\mathbb{Z}) \to V(\mathbb{Z}/p^a\mathbb{Z})$ for some $a \geq 1$, for which we can compute or bound the Fourier transform. (Incorporating the trivial bound $|\widehat{\Psi_p}(x)| \ll 1$ yields results which are in some sense nontrivial, but our interest is in doing better.)

In [28], explicit formulas like (6) are computed for any function Ψ_p for which $\Psi_p(gx) = \Psi_p(x)$ for all $g \in G(\mathbb{F}_p)$ when $a = 1$.

4. (*Orbits in geometric terms.*) For each orbit description $\mathcal{O}$, there exists an integer $i = i(\mathcal{O}) \in [0, d]$ such that $\#\mathcal{O}(p) \asymp_{\mathcal{O}} p^i$ as p ranges; we call this integer the dimension of $\mathcal{O}$. We will show in Section 5 that there also exists a closed *subscheme* $\mathfrak{X} \subseteq V$ defined over $\mathbb{Z}$ of the same dimension $i(\mathcal{O})$, such that $\mathcal{O}(p) \subseteq \mathfrak{X}(\mathbb{F}_p)$ for all but (possibly) finitely many primes p. As we will see, this will allow lattice point counting methods which use algebraic geometry.

Remark 3 *The 'schemes' in question are simply the vanishing loci of systems of polynomials defined over $\mathbb{Z}$, and the algebraic geometry to be invoked will be fairly elementary. However, one can study related problems using very sophisticated algebro-geometric tools; see for example [18, 20].*

2 Levels of distribution and the weighted sieve

We begin by discussing this sieve machinery we will apply. In some (but not complete) generality, a *level of distribution* describes the following. Suppose that $a(n) : \mathbb{Z}^+ \to [0, \infty)$ is a function for which we can prove, for each squarefree integer q (including $q = 1$), an estimate of the shape

$$\sum_{\substack{n < X \\ q | n}} a(n) = \omega(q) CY + E(X, q) \tag{7}$$

for some constant C, multiplicative function $\omega(q)$ satisfying $0 \leq \omega(q) < 1$ for all q, function Y of X, and error term $E(X, q)$. With the setup described in Section 1 we will have $Y = X^{r/d}$ with

$$a(n) = \sum_{\substack{x \in V(\mathbb{Z}) \\ |\operatorname{Disc}(x)| = n}} \phi(xX^{-1/d}). \tag{8}$$

We say that the function $a(n)$ has *level of distribution* α if for any $\epsilon > 0$ we have

$$\sum_{q<X^\alpha} |E(X,q)| \ll_\epsilon Y^{1-\epsilon}, \tag{9}$$

where the sum is over squarefree integers q only. Bounds of the shape (9) are required for essentially all sieve methods, and also in many other analytic number theory techniques.

For our formulation of the *weighted sieve* we will also demand a *one-sided linear sieve inequality*

$$\prod_{w\le p<z} (1-\omega(p))^{-1} \le K\left(\frac{\log z}{\log w}\right) \tag{10}$$

for all $2 \le w < z$ and some fixed constant $K \ge 1$; the product is over primes. A familiar computation (see, for example, [21, (5.34)-(5.37)]) shows that (10) holds if we assume for all prime p that $w(p) < 1$ and that

$$\left| w(p) - \frac{1}{p} \right| < \frac{C}{p^2} \tag{11}$$

for a fixed constant C, on which the constant K of (10) depends. (Conditions such as (10) and (11) are often required in sieve methods, and may appear in a variety of guises.)

The *weighted sieve*, developed principally by Richert [25] and Greaves [22], and described here in the formulation of Friedlander and Iwaniec [21, Theorem 25.1], detects almost prime values of n in the sequence $a(n)$.

Theorem 4 (The weighted sieve [25, 22]) *Assume* (7), (9), *and* (10), *and let* $t \ge \frac{1}{\alpha} + \frac{\log 4}{\log 3} - 1$ *be a positive integer. Then we have*

$$\sum_{\substack{n\le X \\ p|n \Rightarrow p > X^{\alpha/4} \\ \nu(n)\le t}} a(n) \gg \frac{Y}{\log X}, \tag{12}$$

where $\nu(n)$ *denotes the number of prime divisors of* n.

This is one of many sieve methods which establish various consequences from hypotheses of the form (7)-(10). We refer to [21] for a nice overview of many different sieve methods and their applications, and to [2, 3, 7, 9, 10, 11, 13, 20, 26, 27] for applications concerning prehomogeneous vector spaces. The papers [2, 7, 9, 13, 27] sieve rings for maximality, where the analogue of $\omega(q)$ is roughly $1/q^2$; conversely, [11, 26] illustrate sieves where $\omega(q)$ is not a decreasing function of q.

Technical notes. Theorem 4 may be deduced from the precise statement of Theorem 25.1 of [21] as follows. We take $N = 1$ in (25.7), corresponding to (7). We choose $u = 1$ so that $\delta(u) = \frac{\log 4}{\log 3}$ (this is the limit as $u \to 1$ of the expression in (25.17)), and we have $V(X) \gg_\alpha \log X$ by (10).

As mentioned in [21], Greaves proved [22, Chapter 5] a related result with $\frac{\log 4}{\log 3} = 1.261\ldots$ replaced with $1.124\ldots$ Since we will eventually obtain $\alpha = \frac{7}{48}$ for the representation $\mathbb{Z}^2 \otimes \mathrm{Sym}_2(\mathbb{Z}^3)$, this would yield quartic field discriminants with only 7 prime factors in Theorem 2. But since our main goal is to showcase our sieve method, we have chosen to apply a form of the weighted sieve that is easier to extract from the literature. Note that the lower bound $p > X^{\alpha/4}$ will be important in Section 8.

3 Smoothing and the Poisson summation formula

Assumptions. Until Section 3.1, the analysis in this section is quite general (and very standard). $V(\mathbb{Z})$ will denote a complete lattice in a vector space $V(\mathbb{R})$ of dimension r. (V itself will denote an r-dimensional affine space over $\mathbb{Z}$.) In what follows d will be the (homogeneous) degree of the discriminant polynomial, but in

this section (where such a polynomial need not be defined) d may be any positive real constant. X will be a positive real number; and for each squarefree integer q, $\Psi_q : V(\mathbb{Z}) \to \mathbb{C}$ is any function which factors through the reduction map $V(\mathbb{Z}) \to V(\mathbb{Z}/q\mathbb{Z})$. All sums over q will implicitly be over *squarefree* positive integers q only. We assume for simplicity that $|\Psi_q(x)| \leq 1$ for all q and x. Finally, ϕ will denote any fixed smooth, Schwartz class function, on which all implied constants below are allowed to depend.

The aim of this section is to estimate the values of the sum

$$\sum_{x \in V(\mathbb{Z})} \Psi_q(x)\phi(xX^{-1/d}), \tag{13}$$

and in particular to prove upper bounds for the error terms, summed over q. In the general setting of Section 1 this is a smoothed undercount of those $x \in G(\mathbb{Z})\backslash V(\mathbb{Z})$ with $0 < \pm \operatorname{Disc}(x) < X$ satisfying the congruence conditions implied by the function Ψ_q. In the more specific setting of the proof of Theorem 2, Ψ_q is the characteristic function of those x with $q \mid \operatorname{Disc}(x)$, so that (13) counts discriminants divisible by q.

By Poisson summation and a standard unfolding argument, we may check that

$$\sum_{x \in V(\mathbb{Z})} \Psi_q(x)\phi(xX^{-1/d}) = X^{r/d} \sum_{x \in V^*(\mathbb{Z})} \widehat{\Psi_q}(x)\widehat{\phi}\left(\frac{xX^{1/d}}{q}\right) \tag{14}$$

$$= \widehat{\Psi_q}(0)\widehat{\phi}(0)X^{r/d} + E(X, \Psi_q, \phi), \tag{15}$$

where the error term $E(X, \Psi_q, \phi)$ is defined by

$$E(X, \Psi_q, \phi) := X^{r/d} \sum_{0 \neq x \in V^*(\mathbb{Z})} \widehat{\Psi_q}(x)\widehat{\phi}\left(\frac{xX^{1/d}}{q}\right), \tag{16}$$

and $\widehat{\phi}$ satisfies the rapid decay property

$$|\widehat{\phi}(y)| \ll_A (1 + |y|)^{-A} \tag{17}$$

for every $A > 0$ and every $y \in V^*(\mathbb{R})$. (Here $|y|^2 := y_1^2 + \cdots + y_d^2$.)

With an eye to (9), we desire the following conclusion:

Conclusion 5 (Level of distribution α) *We have, for a parameters $\alpha > 0$ to be determined, that the following inequality holds for some $c < r/d$:*

$$\sum_{q < X^\alpha} |E(X, \Psi_q, \phi)| \ll X^c. \tag{18}$$

We will now prove that this conclusion is implied by the more combinatorial statements of (24) in Proposition 7 or (25) in Proposition 8.

For a parameter $Z > 0$, denote by $E_{\leq Z}(X, \Psi_q, \phi)$ the contribution to $E(X, \Psi_q, \phi)$ from those x whose coordinates are all bounded by Z, and write $E_{>Z}(X, \Psi_q, \phi)$ for the remaining contribution.

Lemma 6 *For any $Z > 0$ and $A > d$ we have*

$$E_{>Z}(X, \Psi_q, \phi) \ll_A X^{r/d} \left(\frac{q}{X^{1/d}}\right)^A Z^{-A+d}, \tag{19}$$

and if $Z := qX^{-1/d+\eta}$ for a fixed constant $\eta > 0$ and $q < X$ we have, for any $B > 0$,

$$E_{>Z}(X, \Psi_q, \phi) \ll_{B,\eta} X^{-B}. \tag{20}$$

Proof: By (17), we have

$$E_{>Z}(X,\Psi_q,\phi) \ll_A X^{r/d} \sum_{\substack{x \\ \exists i\ |x_i|>Z}} \left(1+\frac{|x|X^{1/d}}{q}\right)^{-A} \leq X^{r/d}\left(\frac{q}{X^{1/d}}\right)^A \sum_{\substack{x \\ |x|>Z}} |x|^{-A}. \tag{21}$$

There are $\ll R^d$ elements x with $|x| \in [R,2R]$ for any $R>0$. Therefore, assuming that $A>d$ this sum is

$$\ll X^{r/d}\left(\frac{q}{X^{1/d}}\right)^A \sum_{j=0}^{\infty} (2^j Z)^{-A+d} \ll X^{r/d}\left(\frac{q}{X^{1/d}}\right)^A Z^{-A+d}, \tag{22}$$

proving (19). With $Z := qX^{-1/d+\eta}$ this simplifies to $X^{\frac{r}{d}-1+(d-A)\eta}q^d \leq X^{\frac{r}{d}-1+d+(d-A)\eta}$, and the result follows by choosing $A = \frac{B+d+\frac{r}{d}-1}{\eta}+d$. □

In what follows we will choose $Z = Z(q) := qX^{-1/d+\eta}$ for a fixed small $\eta>0$ so as to guarantee (20), so that we have $E(X,\Psi_q,\phi) = O_{B,\eta,\phi}(X^{-B}) + E_{\leq Z}(X,\Psi_q,\phi)$, with

$$|E_{\leq Z}(X,\Psi_q,\phi)| \leq X^{r/d}\widehat{\phi}(0) \sum_{\substack{0\neq x\in V^*(\mathbb{Z}) \\ |x_i|\leq Z\ \forall i}} |\widehat{\Psi_q}(x)|, \tag{23}$$

with $\widehat{\phi}(0)$ being a convenient upper bound for $|\widehat{\phi}(t)|$. We remark that if $q < X^{1/d-\eta}$ then the sum in (23) is empty and $E(X,\Psi_q,\phi) \ll_B X^{-B}$; i.e., the error is essentially zero. In general, we conclude the following:

Proposition 7 (Level of distribution α, simplified version) *Conclusion 5 follows if we have, for the same $\alpha>0$, some $c<r/d$ and $\eta>0$, every $N<X^\alpha$, and with $Z := 2NX^{\eta-1/d}$, that*

$$X^{r/d} \sum_{q\in[N,2N]} \sum_{\substack{0\neq x\in V^*(\mathbb{Z}) \\ |x_i|\leq Z\ \forall i}} |\widehat{\Psi_q}(x)| \ll X^c. \tag{24}$$

Proof: We divide the sum in (18) into $\ll \log X$ dyadic intervals $[N,2N]$ and apply (23) to each $E(X,\Psi_q,\phi)$, for each q expanding the condition $|x_i| \leq Z(q)$ to $|x_i| \leq Z(N) = 2NX^{-1/d+\eta}$. The term $\widehat{\phi}(0)$ may be subsumed (for fixed ϕ) into constants implied by the notation $\ll$ and $O(-)$, and Conclusion 5 follows (with any c strictly larger than that in (24), so as to incorporate a contributions of $O(\log X)$ from the number of intervals). □

This statement may initially look more complicated than Conclusion 5, but it is simpler in that it lends itself naturally to geometric proofs. Moreover the sums over q and x are independent and can be interchanged.

The L_1 norm heuristic. In the introduction, we said that 'L_1 norm bounds for Fourier transforms over finite fields should lead to level of distribution statements for arithmetic objects.' Such a heuristic arises from (24) by assuming that $|\widehat{\Psi_q}(x)|$ has the same average value in the box defined by $|x_i| \leq Z$ as it does in all of $V^*(\mathbb{Z}/q\mathbb{Z})$. Such a statement cannot be proved in general, and indeed it is not always true: for example, in $\mathbb{Z}^2 \otimes \mathrm{Sym}_2\,\mathbb{Z}^3$ there are disproportionately many doubled forms $x=(x_1,x_1) \in V^*(\mathbb{Z}) \cap [-Z,Z]^{12}$ near the origin. That said, this heuristic is the motivation behind our geometric sieve method, and it also provides a target for what we may hope to prove.

3.1 Reformulation in terms of $V(\mathbb{Z})$

In practice it will be convenient to describe the Fourier transforms $\widehat{\Psi_q}(x)$ in terms of $V(\mathbb{Z})$ instead of $V^*(\mathbb{Z})$. To do this, assume we have a linear map $\rho : V^* \to V$, defined by equations over $\mathbb{Z}$, satisfying

the following properties[1] for some integer m: (1) We have $mV(\mathbb{Z}) \subseteq \rho(V^*(\mathbb{Z})) \subseteq V(\mathbb{Z})$; (2) ρ defines an isomorphism $V^*(\mathbb{Z}/q\mathbb{Z}) \to V(\mathbb{Z}/q\mathbb{Z})$ for all integers q coprime to m; (3) for each $x \in V^*(\mathbb{Z})$, the coefficients of $\rho(x) \in V(\mathbb{Z})$ are bounded above by m times those of x. Note that (2) is implied by (1), since $mV(\mathbb{Z}/q\mathbb{Z}) \subseteq \rho(V^*(\mathbb{Z}/q\mathbb{Z})) \subseteq V(\mathbb{Z}/q\mathbb{Z})$ for each q.

We then define $\widehat{\Psi_q}$ on $V(\mathbb{Z}/q\mathbb{Z})$ by writing $\widehat{\Psi_q}(x) = \widehat{\Psi_q}(\rho^{-1}(x))$, and we lift this definition of $\widehat{\Psi_q}$ to all of $V(\mathbb{Z})$. Finally, by abuse of notation we write $\widehat{\Psi_q}(x) = \widehat{\Psi_{\frac{q}{(q,m)}}}(x)$ for an arbitrary squarefree q, so that we have defined $\widehat{\Psi_q}(x)$ for all squarefree q and all $x \in V(\mathbb{Z})$.

The following is then immediate:

Proposition 8 (Level of distribution α, simplified version in terms of $V(\mathbb{Z})$) *Given the constructions above, Conclusion 5 follows if we have, for the same $\alpha > 0$, some $c < r/d$ and $\eta > 0$, every $N < X^\alpha$, and with $Z := NX^{\eta-1/d}$, that*

$$X^{r/d} \sum_{q \in [N,2N]} \sum_{\substack{0 \neq x \in V(\mathbb{Z}) \\ |x_i| \leq Z \ \forall i}} |\widehat{\Psi_q}(x)| \ll X^c. \tag{25}$$

In fact this conclusion is immediate only with $Z = 2mNX^{\eta-1/d}$, but we observe that we may divide all our previous choices of Z by $2m$, with identical results holding at each step; alternatively we may take η larger than that of Proposition 7. The implied constant in (25) is independent of N and X but may depend on the other variables.

For each of the two specific representations (G, V) we treat in this paper, as well as many other cases of interest, such a ρ is naturally induced by a non-degenerate symmetric bilinear form $[-,-]$ on V, defined over $\mathbb{Z}[1/m]$, for which $[gx, g^\iota y] = [x, y]$ identically for an involution ι of G. Whenever $(q, m) = 1$, this implies that $\rho : V^*(\mathbb{Z}/q\mathbb{Z}) \to V(\mathbb{Z}/q\mathbb{Z})$ defines an isomorphism of $G(\mathbb{Z}/q\mathbb{Z})$-modules. These facts are important to our evaluation of the Fourier transforms $\widehat{\Psi_q}$ in [28], which we describe as functions on $V(\mathbb{F}_q)$ rather than on $V^*(\mathbb{F}_q)$.

We refer to [28], especially Sections 2 and A, for further details and explicit constructions. For example, let V be the space of binary cubic forms with $G = \mathrm{GL}_2$. Then $V(\mathbb{Z})$ is the lattice of all integral binary cubic forms and $m = 3$. The bilinear form on V is defined by

$$[x, x'] = aa' + \frac{1}{3}bb' + \frac{1}{3}cc' + dd'$$

for $x = au^3 + bu^2v + cuv^2 + dv^3$ and $x' = a'u^3 + b'u^2v + c'uv^2 + d'v^3$; the involution ι is defined by $g \mapsto g^{-T}$; and ρ is the inverse of the map $V \ni x \mapsto [\cdot, x] \in V^*$, which is an isomorphism over $\mathbb{Z}[1/3]$. Since $V^*(\mathbb{Z}) = \{\phi \in V^*(\mathbb{Q}) \mid \phi(V(\mathbb{Z})) \subset \mathbb{Z}\}$, under the identification $V(\mathbb{Q}) = V^*(\mathbb{Q})$, $V^*(\mathbb{Z})$ is the lattice of integral binary cubic forms whose two middle coefficients are multiples of 3, and thus $V^*(\mathbb{Z}) \subset V(\mathbb{Z})$.

For the space V of pairs of ternary quadratic forms, $V(\mathbb{Z})$ is the lattice of all pairs of integral quadratic forms, $V^*(\mathbb{Z}) \subset V(\mathbb{Z})$ is the lattice of pairs of integral quadratic forms whose off-diagonal coefficients are multiples of 2, and $m = 2$.

4 Proof of Theorem 1

We now prove Theorem 1 by obtaining a level of distribution of $\frac{1}{2} - \epsilon$ for a smoothed subset of integral orbits of binary cubic forms (where the level of distribution is again taken with respect to the property of the discriminant being divisible by q). Although we could appeal to the geometric sieve method of Section 7, we instead give an easier proof whose idea is roughly equivalent to a special case of this method.

[1] Formally we may define ρ as a morphism of schemes over $\mathbb{Z}$ (which is an isomorphism over $\mathbb{Z}[1/m]$); what we need is that ρ defines maps $V^*(\mathbb{Z}) \to V(\mathbb{Z})$, $V^*(R) \to V(R)$ for each ring R containing $\mathbb{Z}$, and $V^*(\mathbb{Z}/q\mathbb{Z}) \to V(\mathbb{Z}/q\mathbb{Z})$ for each quotient $\mathbb{Z}/q\mathbb{Z}$ of $\mathbb{Z}$, all defined by the same equations and hence compatible with the appropriate ring homomorphisms.

In this section $V(\mathbb{Z}) := \mathrm{Sym}_3 \mathbb{Z}^2$ is the lattice of integral binary cubic forms, $r = d = 4$, $Z := NX^{-1/4+\eta}$, Ψ_q is the characteristic function of $x \in V(\mathbb{Z})$ with $q \mid \mathrm{Disc}(x)$, and we argue that we can prove the bound (25) for any $\alpha < \frac{1}{2}$, i.e. that

$$\sum_{q\in[N,2N]} \sum_{\substack{0\neq x\in V(\mathbb{Z}) \\ |x_i|\leq Z \ \forall i}} |\widehat{\Psi_q}(x)| \ll X^{-1+c} \tag{26}$$

for each $N < X^\alpha$ for some $c = c(\alpha) < 1$. The Fourier transform $\widehat{\Psi_q}$ is multiplicative in q, and satisfies

$$\widehat{\Psi_p}(x) = \begin{cases} p^{-1} + p^{-2} - p^{-3} & \text{if } x \in pV(\mathbb{Z}), \\ p^{-2} - p^{-3} & \text{if } x \notin pV(\mathbb{Z}) \text{ but } p \mid \mathrm{Disc}(x), \\ -p^{-3} & \text{if } p \nmid \mathrm{Disc}(x). \end{cases} \tag{27}$$

For each positive divisor q_0 of q and $x \in q_0 V(\mathbb{Z})$, we have

$$\widehat{\Psi_q}(x) = \widehat{\Psi_{q_0}}(x) \cdot \widehat{\Psi_{q/q_0}}(x) = \widehat{\Psi_{q_0}}(x) \cdot \widehat{\Psi_{q/q_0}}(x/q_0).$$

Therefore the summation of (26) is equal to

$$\sum_{q_0\leq Z} \Big(\prod_{p\mid q_0} (p^{-1} + p^{-2} - p^{-3}) \Big) \sum_{\substack{\frac{N}{q_0} \leq q_1 \leq \frac{2N}{q_0} \\ (q_0,q_1)=1}} \sum_{\substack{x \\ |x_i|\leq \frac{Z}{q_0} \ \forall i}}{}' |\widehat{\Psi_{q_1}}(x)|, \tag{28}$$

where the inner sum is over those $x \in V(\mathbb{Z})$ which are not in $pV(\mathbb{Z})$ for any prime divisor p of $q_0 q_1$.

We split the sum of (28) into two pieces: a sum over those x for which $\mathrm{Disc}(x) = 0$, and a sum over those x for which $\mathrm{Disc}(x) \neq 0$.

Those x with $\mathrm{Disc}(x) = 0$. The number of such x with all coordinates bounded by Y, is $O(Y^2)$ for any Y. To see this, note that any such x can be written as $(ax+by)^2(cx+dy)$ for some $a, b, c, d \in \mathbb{Z}$. The number of possibilities with $a = 0$ is $O(Y^2)$, as this forces the x^3 and x^2y coefficients to both be zero. Similarly there are $O(Y^2)$ possibilities with $b = 0$. We are therefore left with the number of integer quadruples (a, b, c, d) with $ab \neq 0$, $|a^2c| \leq Y$, and $|b^2d| \leq Y$, which is the square of the number of integer pairs (a, c) with $a \neq 0$, $|a^2c| \leq Y$. This latter quantity is easily seen to be $O(Y)$, as needed.

The inner sum is therefore over $O(Z/q_0)^2$ elements, and for each x we have $|\widehat{\Psi_{q_1}}(x)| \leq {q_1}^{-2}$. Therefore, this portion of the sum in (28) is

$$\ll_\epsilon \sum_{q_0\leq Z} q_0^{-1+\epsilon} \cdot \frac{N}{q_0} \cdot \left(\frac{Z}{q_0}\right)^2 \cdot \left(\frac{N}{q_0}\right)^{-2} \ll_\epsilon N^{-1} Z^2 \sum_{q_0\leq Z} q_0^{-2+\epsilon} \ll X^{2\eta} N X^{-1/2}, \tag{29}$$

which satisfies the bound (26).

Those x with $\mathrm{Disc}(x) \neq 0$. The contribution of these is bounded above by

$$\begin{aligned} &\sum_{q_0\leq Z} \Big(\prod_{p\mid q_0} (p^{-1} + p^{-2} - p^{-3}) \Big) \sum_{\substack{\frac{N}{q_0} \leq q_1 \leq \frac{2N}{q_0} \\ (q_0,q_1)=1}} q_1^{-3} \sum_{\substack{x \\ |x_i|\leq \frac{Z}{q_0} \ \forall i}}{}' \gcd(\mathrm{Disc}(x), q_1) \\ &\leq \sum_{q_0\leq Z} \Big(\prod_{p\mid q_0} (p^{-1} + p^{-2} - p^{-3}) \Big) \left(\frac{N}{q_0}\right)^{-3} \sum_{\substack{x \\ |x_i|\leq Z/q_0 \ \forall i \\ \mathrm{Disc}(x)\neq 0}} \sum_{\frac{N}{q_0} \leq q_1 \leq \frac{2N}{q_0}} \gcd(\mathrm{Disc}(x), q_1). \end{aligned}$$

Now, in general, whenever $m \neq 0$ we have

$$\sum_{n\in[N,2N]} \gcd(m,n) \leq \sum_{\substack{f|m\\ f\leq 2N}} f \sum_{\substack{n\in[N,2N]\\ f|n}} 1 \leq \sum_{\substack{f|m\\ f\leq 2N}} f\left(\frac{N}{f}+1\right) \ll Nm^{\epsilon}.$$

Therefore, using that the discriminant of any x in the sum is $\ll N^4$, we see that the previous quantity is

$$\begin{aligned}
&\ll_{\epsilon} \sum_{q_0\leq Z} \Big(\prod_{p|q_0}(p^{-1}+p^{-2}-p^{-3})\Big)\left(\frac{N}{q_0}\right)^{-3}\cdot\left(\frac{Z}{q_0}\right)^4\frac{N}{q_0}\cdot N^{\epsilon}\\
&\ll_{\epsilon} N^{\epsilon}Z^4N^{-2}\sum_{q_0\leq Z} q_0^{-3+\epsilon}\\
&\ll_{\epsilon} N^{\epsilon}N^2X^{-1+4\eta},
\end{aligned}$$

again satisfying (26).

Applying the weighted sieve of Theorem 4, and following the beginning of the proof of Proposition 12, we obtain $\gg \frac{X}{\log X}$ elements $x \in V(\mathbb{Z})$ whose discriminants have at most three prime factors. Of these, at most $O_{\epsilon}(X^{3/4+\epsilon})$ can be reducible. (For a simple proof see [13, Lemma 21]; only the second paragraph of the proof there is relevant, as we are counting points in a box of side length $\ll X^{1/4}$.) As the weighted sieve produces x with each prime factor $> X^{\alpha/4}$, the number of $x \in V(\mathbb{Z})$ with any repeated prime factor is $\ll \sum_{p>X^{\alpha/4}} \frac{X}{p^2} \ll X^{1-\alpha/4}$ by [2, Lemma 3.4]. Accordingly we produce $\gg \frac{X}{\log X}$ irreducible elements $x \in V(\mathbb{Z})$ with squarefree (and hence fundamental) discriminants, which must therefore correspond to (distinct) maximal cubic orders and hence to cubic fields.

Remark 9 *In place of our estimate of $O(Y^2)$ for reducible elements x in boxes of side length $O(Y)$, the method of Section 7 would implicitly incorporate a bound of $O(Y^3)$, as 3 is the dimension of the variety* $\mathrm{Disc}(x) = 0$. *This proof illustrates that counting elements more directly may yield improvements in the end results.*

5 Closed subschemes containing singular orbits

Let V be the space of pairs of ternary quadratic forms, together with its action of $G = \mathrm{GL}_2 \times \mathrm{GL}_3$. Recall from [28, Proposition 21] that, for each $p \neq 2$ there are 20 orbits for the action of $G(\mathbb{F}_p)$ on $V(\mathbb{F}_p)$. We gave twenty 'orbit descriptions' $\mathcal{O}$ which were essentially uniform in p, and for each p we write $\mathcal{O}(p)$ for the associated orbit over $\mathbb{F}_p$.

Proposition 10 *For each of the orbit descriptions $\mathcal{O}$ described above there exists a closed subscheme $\mathfrak{X} \subset V$, defined over $\mathbb{Z}$, such that $\mathcal{O}(p) \subseteq \mathfrak{X}(\mathbb{F}_p)$ for each prime $p \neq 2$, and of the same 'dimension' as $\mathcal{O}$ in the sense that $\#\mathcal{O}(p) \asymp \#\mathfrak{X}(\mathbb{F}_p)$.*

We will prove this statement, with '$p \neq 2$' replaced with '$p \notin S$ for some finite set S', for any finite dimensional (G, V) satisfying the following two properties:

- There exist finitely many elements $x_\sigma \in V(\mathbb{Z})$ such that for any algebraically closed field K with $\mathrm{char}(K) \notin S$, the images of the x_σ in $V(K)$ via the canonical map $V(\mathbb{Z}) \to V(K)$ form a set of complete representatives for $G(K)\backslash V(K)$.
- There exists a constant $c > 0$ such that for each $p \notin S$ and $G(\mathbb{F}_p)$-orbit $\mathcal{O} \subseteq V(\mathbb{F}_p)$ we have $\#\mathcal{O} > c\#\widetilde{\mathcal{O}}$, with $\widetilde{\mathcal{O}} := G(\overline{\mathbb{F}_p})\mathcal{O} \cap V(\mathbb{F}_p)$.

These properties hold for all of the (G, V) studied in [28], as we explain now in the case of pairs of ternary quadratic forms. We group the 20 orbit descriptions of [28, Proposition 21] as follows:

$$\{\mathcal{O}_0\}, \{\mathcal{O}_{D1^2}\}, \{\mathcal{O}_{D11}, \mathcal{O}_{D2}\}, \{\mathcal{O}_{D\mathrm{ns}}\}, \{\mathcal{O}_{C\mathrm{s}}\}, \{\mathcal{O}_{C\mathrm{ns}}\}, \{\mathcal{O}_{B11}, \mathcal{O}_{B2}\},$$
$$\{\mathcal{O}_{1^4}\}, \{\mathcal{O}_{1^31}\}, \{\mathcal{O}_{1^21^2}, \mathcal{O}_{2^2}\}, \{\mathcal{O}_{1^211}, \mathcal{O}_{1^22}\}, \{\mathcal{O}_{1111}, \mathcal{O}_{112,}\mathcal{O}_{22}, \mathcal{O}_{13}, \mathcal{O}_4\}.$$

Within each of these twelve sets, the orbital representative of the *first-listed* $\mathcal{O}$ is described in [28, Proposition 21] as the reduction $\pmod p$ of a fixed element $x_\sigma \in V(\mathbb{Z})$, and when K is algebraically closed with $\mathrm{char}(K) \neq 2$, the proof in [28, Section 7.1] establishes that the images of these x_σ in $V(K)$ are a set of representatives for $G(K)\backslash V(K)$. Moreover, for each $p \notin S$ and $G(\mathbb{F}_p)$-orbit $\mathcal{O} \subseteq V(\mathbb{F}_p)$, the associated $\widetilde{\mathcal{O}}$ is the union of the $\mathcal{O}$ in the grouping described above, and the second property above may be deduced from the point counts in [28, Proposition 21].

To conclude Proposition 10 from these two properties, write $Y_\sigma := G(\overline{\mathbb{Q}})x_\sigma \subseteq V(\overline{\mathbb{Q}})$ for each x_σ. By [14, Propositions I.1.8 and II.6.7] we may write each Y_σ in the form $Y_\sigma = \mathfrak{X}_\sigma - \cup_j W_{\sigma,j}$ where the $\mathfrak{X}_\sigma$ and $W_{\sigma,j}$ are (finitely many) closed varieties, defined over $\mathbb{Q}$, and with $\dim(W_{\sigma,j}) < \dim(\mathfrak{X}_\sigma)$ for all σ and j. (Each $\mathfrak{X}_\sigma$ is the closure of Y_σ, and the $W_{\sigma,j}$ are defined by the closures of other $G(\overline{\mathbb{Q}})$-orbits, of which there are finitely many, and since each of the x_σ is defined over $\mathbb{Q}$ their orbits are as well.)

We choose (arbitrary) integral models for the $\mathfrak{X}_\sigma$ so as to regard them as closed subschemes of V. For all but finitely many p, these equations reduce $\pmod p$ and define varieties of the same dimension over $\mathbb{F}_p$, and we conclude by Lang-Weil [23] that $\#\mathfrak{X}_\sigma(\mathbb{F}_p) \asymp \#Y_\sigma(\mathbb{F}_p) \asymp p^{\dim(\mathfrak{X}_\sigma)}$. The second bullet point above then gives the desired conclusion.

6 A version of the geometric sieve

The *Ekedahl-Bhargava geometric sieve*, in the form of [10, Theorem 3.3], asserts the following. Let B be a compact region in $\mathbb{R}^r$, let $\mathfrak{X}$ be a closed subscheme of $\mathbb{A}^r_{\mathbb{Z}}$ of codimension $a \geq 1$, and let λ and P be positive real numbers. Then, we have

$$\#\{x \in \lambda B \cap \mathbb{Z}^r \mid x \pmod p \in \mathfrak{X}(\mathbb{F}_p) \text{ for some prime } p > P\} = O_{B,\mathfrak{X}}\left(\frac{\lambda^r}{P^{a-1}\log P} + \lambda^{r-a+1}\right). \tag{30}$$

We introduce a variation with two modifications. Firstly, we count each x with multiplicity, given by the number of pairs (x, p) for which $x \pmod p \in \mathfrak{X}(\mathbb{F}_p)$ and $p \in [P, 2P]$. Secondly, we introduce an 'arithmetic progression' condition $x - x_0 \in mV(\mathbb{Z})$, allowing for additional flexibility in applications (as we will see in Section 7.)

We refer also to [11] where the same generalization is presented concurrently; the application there replaces *primes* in (30) with squarefree integers, which amounts roughly to a simpler version of the argument in Section 7 here.

Theorem 11 *Let B, $\mathfrak{X}$, a, and λ be as in the statement of* (30)*, let m be a positive integer, let $x_0 \in V(\mathbb{Z})$, and let $P > \lambda/m > 1$ be any real number. Then, we have*

$$\#\{(x,p) \mid x \in \lambda B \cap (x_0 + m\mathbb{Z}^r),\ \ p \text{ is a prime in } [P,2P],\ p \nmid m,\ x \pmod p \in \mathfrak{X}(\mathbb{F}_p)\} \ll_{B,\mathfrak{X},\epsilon} \left(\frac{\lambda}{m}\right)^{r-a} P\lambda^\epsilon. \tag{31}$$

Proof: This closely follows the proof of Theorems 3.1 and 3.3 of Bhargava [10]. In [10] (with $m = 1$), the quantity λ appears only as an upper bound for the number of lattice points in rB lying on a line defined by fixing all but one of the coordinates. Therefore, with $m > 1$ we can replace λ with $\frac{\lambda}{m}$ at each occurrence.

The analogue of [10, Lemma 3.1], proved identically, thus reads that

$$\#\{x \in \lambda B \cap (x_0 + m\mathbb{Z}^r) \cap \mathfrak{X}(\mathbb{Z})\} \ll_{B,\mathfrak{X}} \left(\frac{\lambda}{m}\right)^{r-a}, \tag{32}$$

and we obtain the bound of (31) for those $x \in \mathfrak{X}(\mathbb{Z})$.

For those $x \notin \mathfrak{X}(\mathbb{Z})$, it suffices to prove that

$$\#\left\{(x,p) \;\middle|\; \begin{array}{l} x \in \lambda B \cap (x_0 + m\mathbb{Z}^r),\ x \notin \mathfrak{X}(\mathbb{Z}), \\ p > \frac{\lambda}{m},\ p \nmid m,\ x \pmod p \in \mathfrak{X}(\mathbb{F}_p) \end{array}\right\} \ll_{B,\mathfrak{X},\epsilon} \left(\frac{\lambda}{m}\right)^{r-a+1} \cdot \lambda^{\epsilon}, \tag{33}$$

the exact analogue of [10, (17)]. This follows [10] exactly. The condition $p \nmid m$ is needed in the last paragraph of [10, Theorem 3.3], to conclude that if $f_k(x)$ is a polynomial in one variable with $f_k \not\equiv 0 \pmod p$, then it has $O_{\deg(f_k)}(1)$ roots x in an interval of length $O(\lambda)$, and with $x \equiv x_0 \pmod m$ for any fixed x_0. The factor of λ^{ϵ} arises in adapting the argument immediately after [10, (17)]: any nonzero $f_i(x)$ can have only $O_{f_i}(1)$ prime factors $p > \lambda$, but it may have $O_{f_i,\epsilon}(\lambda^{\epsilon})$ prime factors $p > \lambda/m$. □

7 Application of the geometric sieve: Proof of Proposition 8

In this section we prove the bound (25) for each $\alpha < \frac{7}{48}$ for the 'quartic' (G, V) of Section 1.

For each $p \neq 2$ and $i \in \{0, 4, 7, 8, 10, 11, 12\}$ we define sets $U_i(p)$, each of which is a union of $G(\mathbb{F}_p)$-orbits on $V(\mathbb{F}_p)$, as follows.

Label	Consists of	Dimension i	Fourier contribution fc(i)
U_0	$\mathcal{O}_0$	0	-1
U_4	$\mathcal{O}_{D1^2}$	4	-3
U_7	$\mathcal{O}_{D11}, \mathcal{O}_{C\mathrm{s}}$	7	-4
U_8	$\mathcal{O}_{T11}, \mathcal{O}_{T2}, \mathcal{O}_{D2}, \mathcal{O}_{D\mathrm{ns}}, \mathcal{O}_{C\mathrm{ns}}$	8	-5
U_{10}	$\mathcal{O}_{1^2 1^2}, \mathcal{O}_{2^2}, \mathcal{O}_{1^3 1}, \mathcal{O}_{1^4}$	10	-6
U_{11}	$\mathcal{O}_{1^2 11}, \mathcal{O}_{1^2 2}$	11	-7
U_{12}	nonsingular orbits	12	-8

In Section 5 we proved that for each $i \in \{0, 4, 7, 8, 10, 11, 12\}$ there are subschemes $\mathfrak{X}_i$ of $\mathbb{A}^r_{\mathbb{Z}} = V$ of dimension i, defined over $\mathbb{Z}$, for which $U_i(p) \subseteq \mathfrak{X}_i(\mathbb{F}_p)$ and $\#\mathfrak{X}_i(\mathbb{F}_p) \ll p^i$ for all $p \notin S$. The function fc(i) is chosen such that $|\widehat{\Psi_p}(x)| \leq 2p^{\mathrm{fc}(i)}$ for each $x \in U_i(p)$.

For every squarefree $n \in [N, 2N]$ (with $N < X^{\alpha}$ for α to be determined) we consider the contribution to (25) from every factorization

$$n = n_0 n_4 n_7 n_8 n_{10} n_{11} n_{12} \tag{34}$$

and those x with $x \in U_i(p)$ for each $p \mid n_i$. When n is even we will assume as a bookkeeping device that n_0 is as well, but we will never demand any geometric condition on x modulo 2.

The contribution of each such x is bounded above by $2^{\omega(n)+1} \prod_i n_i^{\mathrm{fc}(i)}$, where the 2^1 factor reflects the trivial bound $|\widehat{\Psi_2}(x)| \leq 1$, and we write $2^{\omega(n)} = O_{\epsilon}(X^{\epsilon})$, uniformly in n.

We consider the following choices of parameters:

- Squarefree and pairwise coprime integers n_i for $i \in \{0, 4, 7, 8, 10, 11, 12\}$, with $\prod_i n_i \in [N, 2N]$.
- A parameter $j \in \{4, 7, 8, 10, 11, 12\}$, and a factorization $n_j = n_j' p n_j''$, where p is a prime.
- Writing $m := n_j' \prod_{i<j} n_i$, these choices are subject to the condition that $m \leq Z < mp$.

We claim that every factorization (34) corresponds to at least one choice of the above data, with $n_j = n_j' p n_j''$. First of all, note that $n_0 \leq Z$ for each nonzero $x \in [-Z, Z]^r$. Thus, given any factorization (34), we let $j \geq 4$ be the minimal index with $\prod_{i \leq j} n_i > Z$, choose n_j' to be the largest divisor of n_j less than or equal to $Z \prod_{i<j} n_i^{-1}$, and choose p to be any prime divisor of n_j / n_j'.

The conditions modulo m, namely that $x \in U_i(q)$ for each odd prime $q \mid m$ with $i = i(q)$ determined by the factorization above, are equivalent to demanding that x lie in one of $O_\epsilon(X^\epsilon n_j'^j \prod_{i<j} n_i^i)$ residue classes (mod $mV(\mathbb{Z})$). (Here X^ϵ is a simple upper bound for $C^{\omega(n)}$, the product of the implied constants occurring in the point counts for the $U_i(p)$.)

We must have $x \in \mathfrak{X}_j(\mathbb{F}_p)$, and for each of the residue classes (mod $mV(\mathbb{Z})$) determined above we use Bhargava's geometric sieve (Theorem 11) to bound the number of pairs (x, p) where $x \in V(\mathbb{Z})$ lies in this residue class, has all coefficients bounded by Z, and lies in $\mathfrak{X}_j(\mathbb{F}_p)$, and where $p \notin S$ lies in a dyadic interval $[P, 2P]$. By the theorem, the number of such pairs is $\ll Z^\epsilon (Z/m)^j P$.

(Any contribution of (x, p) with p in the exceptional set S of Proposition 10 trivially satisfies the same bound, as in this case $Z/m \ll_S 1$.)

The Fourier contribution of each x being counted is $\ll X^\epsilon \prod_i n_i^{\mathrm{fc}(i)}$, and for each choice of j, n_i $(i < j)$, and n_j', and for each fixed dyadic interval $[P, 2P]$, we multiply: the number of residue classes modulo $mV(\mathbb{Z})$; the number of pairs (x, p) in each; the Fourier contribution of each x being counted; and the $N^{1+\epsilon}/mP$ choices of n_j'' and n_i $(i > j)$. Recalling that $Z = NX^{-1/d+\eta}$, we conclude that the contribution to (25) from the choices previously determined is

$$\ll_\epsilon X^\epsilon \cdot X \cdot n_j'^j \prod_{i<j} n_i^i \cdot \left(\frac{NX^{-1/d+\eta}}{m}\right)^j P \cdot \prod_i n_i^{\mathrm{fc}(i)} \cdot \frac{N^{1+\epsilon}}{mP}.$$

Using the fact that $\mathrm{fc}(i)$ is a decreasing function of i, and summing over the $\ll X^\epsilon$ choices of dyadic interval $[P, 2P]$, we see that this is

$$\ll_\epsilon X^{\epsilon+r\eta} \cdot X^{1-j/d} \cdot \left(\prod_{i<j} n_i^{i+\mathrm{fc}(i)}\right) \cdot (n_j')^{j+\mathrm{fc}(j)} \cdot \left(\frac{N}{m}\right)^{j+\mathrm{fc}(j)+1}$$

Now, since $i + \mathrm{fc}(i)$ is an increasing function of i this is bounded above by

$$\begin{aligned} &\ll_\epsilon X^{\epsilon+r\eta} \cdot X^{1-j/d} \cdot m^{j+\mathrm{fc}(j)} \cdot \left(\frac{N}{m}\right)^{j+\mathrm{fc}(j)+1} \\ &\ll_\epsilon X^{\epsilon+r\eta} \cdot X^{1-j/d} \cdot m^{-1} N^{j+\mathrm{fc}(j)+1}, \end{aligned}$$

and, now fixing only the parameter j, we sum over all $m \le Z$ and (for each m) the $\ll N^\epsilon$ choices of factorizations of m to obtain a total contribution

$$\ll_\epsilon X^{\epsilon+r\eta} \cdot X^{1-j/d} \cdot N^{j+\mathrm{fc}(j)+1} \tag{35}$$

from all choices of (34) associated to this factor j. Up to an implied constant, the total error is bounded above by the maximum of (35) over the six admissible values of j. The quantity in (35) is:

j	Bound$(\times X^{\epsilon+r\eta})$
$j = 4$	$X^{2/3}N^2$
$j = 7$	$X^{5/12}N^4$
$j = 8$	$X^{1/3}N^4$
$j = 10$	$X^{1/6}N^5$
$j = 11$	$X^{1/12}N^5$
$j = 12$	N^5

The case $j = 7$ turns out to be the bottleneck, and choosing $N = X^\alpha$ with any $\alpha < \frac{7}{48}$ we may choose η and ϵ with $c := \frac{5}{12} + 4\alpha + \epsilon + 12\eta < 1$, so that (25) holds with this value of c.

8 Conclusion of the proof of Theorem 2

We give a slightly more general statement, which illustrates how improvements to Theorem 2 would automatically follow from improvements in the level of distribution.

Proposition 12 *Assume, for some integer $t \geq 1$, that Proposition 8 (and hence Conclusion 5) holds for some $c < 1$ and $\alpha > \left(t + 1 - \frac{\log 4}{\log 3}\right)^{-1}$. Then there are $\gg_{t,\alpha,c} \frac{X}{\log X}$ S_4-quartic field discriminants K with $|\operatorname{Disc}(K)| < X$, such that $\operatorname{Disc}(K)$ has at most t prime factors.*

Since we proved Proposition 8 with any $\alpha < \frac{7}{48}$, we thus obtain Theorem 2 with any $t > \frac{48}{7} - 1 + \frac{\log 4}{\log 3} = 7.119\ldots$, and in particular with $t = 8$.

Proof: We apply the weighted sieve of Theorem 4, with $Y = X$ and

$$a(n) := \sum_{\substack{x \in V(\mathbb{Z}) \\ |\operatorname{Disc}(x)| = n}} \phi(xX^{-1/12}). \tag{36}$$

Each sum is finite because ϕ is compactly supported. We then have

$$\sum_{\substack{n < X \\ q \mid n}} a(n) = \sum_{x \in V(\mathbb{Z})} \Psi_q(x)\phi(xX^{-1/12}),$$

where Ψ_q is the characteristic function of $x \in V(\mathbb{Z})$ with $q \mid \operatorname{Disc}(x)$. By (14)-(16) the sequence satisfies the sieve axiom (7), and by assumption Proposition 8 and therefore Conclusion 5 and (9) hold. The linearity conditions (10) and (11) follow from the first line of (6).

Theorem 4 therefore implies that the sum of $\phi(xX^{-1/12})$, over all x whose discriminants have at most t prime factors, is $\gg \frac{X}{\log X}$. By construction the count of such x satisfies the same lower bound, and these discriminants are all in $(-X, 0) \cup (0, X)$ and are $G(\mathbb{Z})$-inequivalent in $V(\mathbb{Z})$.

By Bhargava [6, Theorem 1], these are in bijection with pairs (Q, R), where Q is a quartic ring and R is a cubic resolvent ring of R, and in case Q is maximal then it has exactly one cubic resolvent [6, Corollary 5]. Moreover, if $x \in V(\mathbb{Z})$ corresponds to (Q, R), then $\operatorname{Disc}(x) = \operatorname{Disc}(Q)$. As described on [6, p. 1037], Q is an order in an S_4- or A_4-field if and only if the corresponding $x \in V(\mathbb{Z})$ is absolutely irreducible.

The number of x which are not absolutely irreducible is $\ll X^{11/12+\epsilon}$ and hence negligible; this is proved in [6, Lemmas 12 and 13]. In our case these proofs simplify because we may ignore the cusp: the compact support of ϕ ensures that we are only counting points in a box of side length $O(X^{1/12})$, and that the number of points with $a_{11} = 0$ is $O(X^{11/12})$.

We must then bound the number of pairs (Q, R) where Q is a nonmaximal S_4- or A_4-quartic order. By Theorem 4 the discriminants of $x \in V(\mathbb{Z})$ being counted have all of their prime factors $> X^{\alpha/4}$, and in particular any nonmaximal Q which survives the sieve must be nonmaximal at some prime $p > X^{\alpha/4}$. By [7, Proposition 23], the number of such x is

$$\ll \sum_{p > X^{\alpha/4}} X/p^2 \ll X^{1-\alpha/4},$$

negligible for any $\alpha > 0$.

This leaves the maximal Q whose discriminants are divisible by p^2 for some p in the same range. These can be handled by the geometric sieve, precisely as Bhargava did in [10]. We apply the geometric sieve in its original formulation directly to (36), in contrast to Section 7 where we applied our variation after an application of Poisson summation.

Any maximal Q whose discriminant is divisible by p^2 must be, in the language of Bhargava [10], a *strong multiple* of p; and hence (as in Section 5) the corresponding x must be in $\mathfrak{X}(\mathbb{F}_p)$ for a suitably defined

subscheme $\mathfrak{X} \subseteq V$ of codimension 2. By [10, Theorem 3.3], the number of such x which satisfy this criterion for any $p > X^{\alpha/4}$ is again $\ll X^{1-\alpha/4}$.

In conclusion, the contributions of everything other than maximal orders in S_4-quartic fields to our sieve result is negligible, and hence we obtain $\gg \frac{X}{\log X}$ S_4-quartic fields with the stated properties. □

Acknowledgments

We would like to thank Theresa Anderson, Manjul Bhargava, Alex Duncan, Étienne Fouvry, Yasuhiro Ishitsuka, Kentaro Mitsui, Arul Shankar, Ari Shnidman, Jack Thorne, Jerry (Xiaoheng) Wang, Melanie Matchett Wood and Kota Yoshioka for helpful discussions, comments and feedback. Duncan, in particular, explained to us the application of the lemma from [14] to the proof of Proposition 10.

Taniguchi's work was partially supported by the JSPS, KAKENHI Grant Numbers JP24654005, JP25707002, and JP17H02835. Thorne's work was partially supported by the National Science Foundation under Grant No. DMS-1201330 and by the National Security Agency under a Young Investigator Grant.

References

[1] Karim Belabas. Crible et 3-rang des corps quadratiques. *Ann. Inst. Fourier (Grenoble)*, 46(4):909–949, 1996.

[2] Karim Belabas, Manjul Bhargava, and Carl Pomerance. Error estimates for the Davenport-Heilbronn theorems. *Duke Math. J.*, 153(1):173–210, 2010.

[3] Karim Belabas and Etienne Fouvry. Sur le 3-rang des corps quadratiques de discriminant premier ou presque premier. *Duke Math. J.*, 98(2):217–268, 1999.

[4] Manjul Bhargava. Higher composition laws. I. A new view on Gauss composition, and quadratic generalizations. *Ann. of Math. (2)*, 159(1):217–250, 2004.

[5] Manjul Bhargava. Higher composition laws. II. On cubic analogues of Gauss composition. *Ann. of Math. (2)*, 159(2):865–886, 2004.

[6] Manjul Bhargava. Higher composition laws. III. The parametrization of quartic rings. *Ann. of Math. (2)*, 159(3):1329–1360, 2004.

[7] Manjul Bhargava. The density of discriminants of quartic rings and fields. *Ann. of Math. (2)*, 162(2):1031–1063, 2005.

[8] Manjul Bhargava. Higher composition laws. IV. The parametrization of quintic rings. *Ann. of Math. (2)*, 167(1):53–94, 2008.

[9] Manjul Bhargava. The density of discriminants of quintic rings and fields. *Ann. of Math. (2)*, 172(3):1559–1591, 2010.

[10] Manjul Bhargava. The geometric sieve and the density of squarefree values of invariant polynomials. *Preprint*, 2014. Available at `https://arxiv.org/abs/1402.0031`.

[11] Manjul Bhargava, Alina Carmen Cojocaru, and Frank Thorne. The square sieve and the number of A_5-quintic extensions of bounded discriminant. *In preparation.*

[12] Manjul Bhargava and Wei Ho. Coregular spaces and genus one curves. *Cambridge J. of Math.*, 4(1):1–119, 2016.

[13] Manjul Bhargava, Arul Shankar, and Jacob Tsimerman. On the Davenport-Heilbronn theorems and second order terms. *Invent. Math.*, 193(2):439–499, 2013.

[14] Armand Borel. *Linear algebraic groups*, volume 126 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[15] Henri Cohen and Frank Thorne. Dirichlet series associated to quartic fields with given cubic resolvent. *Res. Number Theory*, 2:Art. 29, 40, 2016.

[16] Alina Carmen Cojocaru and M. Ram Murty. *An introduction to sieve methods and their applications*, volume 66 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2006.

[17] H. Davenport and H. Heilbronn. On the density of discriminants of cubic fields. II. *Proc. Roy. Soc. London Ser. A*, 322(1551):405–420, 1971.

[18] Jan Denef and Akihiko Gyoja. Character sums associated to prehomogeneous vector spaces. *Compositio Math.*, 113(3):273–346, 1998.

[19] Jordan Ellenberg, Lillian B. Pierce, and Melanie Matchett Wood. On ℓ-torsion in class groups of number fields. *Preprint*, 2016. Available at `https://arxiv.org/abs/1606.06103`.

[20] E. Fouvry and N. Katz. A general stratification theorem for exponential sums, and applications. *J. Reine Angew. Math.*, 540:115–166, 2001.

[21] John Friedlander and Henryk Iwaniec. *Opera de cribro*, volume 57 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2010.

[22] George Greaves. *Sieves in number theory*, volume 43 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 2001.

[23] Serge Lang and André Weil. Number of points of varieties in finite fields. *Amer. J. Math.*, 76:819–827, 1954.

[24] L. B. Pierce and F. Thorne. *Voronoi summation formulae related to the Shintani zeta function and applications to twisted averages*. In preparation.

[25] H.-E. Richert. Selberg's sieve with weights. *Mathematika*, 16:1–22, 1969.

[26] Arul Shankar and Jacob Tsimerman. Counting S_5-fields with a power saving error term. *Forum Math. Sigma*, 2:e13, 8, 2014.

[27] Takashi Taniguchi and Frank Thorne. Secondary terms in counting functions for cubic fields. *Duke Math. J.*, 162(13):2451–2508, 2013.

[28] Takashi Taniguchi and Frank Thorne. Orbital exponential sums for prehomogeneous vector spaces. *Preprint*, 2016. Available at `https://arxiv.org/abs/1607.07827`.

[29] David J. Wright and Akihiko Yukie. Prehomogeneous vector spaces and field extensions. *Invent. Math.*, 110(2):283–314, 1992.