Squarefree values of polynomial discriminants II

Manjul Bhargava, Arul Shankar, and Xiaoheng Wang

July 13, 2022

Abstract

We determine the density of integral binary forms of given degree that have squarefree discriminant, proving for the first time that the lower density is positive. Furthermore, we determine the density of integral binary forms that cut out maximal orders in number fields. The latter proves, in particular, an “arithmetic Bertini theorem” conjectured by Poonen for $\mathbb{P}^1_\mathbb{Z}$.

Our methods also allow us to prove that there are $\gg X^{1/2+1/(n-1)}$ number fields of degree $n$ having associated Galois group $S_n$ and absolute discriminant less than $X$, improving the best previously known lower bound of $\gg X^{1/2+1/n}$.

Finally, our methods correct an error in and thus resurrect earlier (retracted) results of Nakagawa on lower bounds for the number of totally unramified $A_n$-extensions of quadratic number fields of bounded discriminant.

1 Introduction

In the first article [10] of this two-part series, we proved that when monic integer polynomials $f(x) = x^n + a_1 x^{n-1} + \cdots + a_n$ of fixed degree $n$ are ordered by $\max\{|a_1|, \ldots, |a_n|^{1/n}\}$, a positive proportion have squarefree discriminant. The purpose of this article is to prove the analogous result for integral binary $n$-ic forms.

Recall that the discriminant $\Delta(f)$ of a binary $n$-ic form over a field $K$ is a homogeneous polynomial of degree $2n-2$ in the coefficients of $f$, whose nonvanishing is equivalent to $f$ having $n$ distinct linear factors over an algebraic closure $\overline{K}$ of $K$. We order integral binary $n$-ic forms $f(x,y) = a_0 x^n + a_1 x^{n-1} y + \cdots + a_n y^n$ by their height $H(f)$ given by $H(f) := \max\{|a_0|, \ldots, |a_n|\}$, i.e., the maximum of the absolute values of the coefficients. Then a natural question is: when ordered by height, what is the density of integral binary $n$-ic forms whose discriminant is squarefree? For $n=2$, classical methods in sieve theory yield the answer. For $n=3$ and $n=4$, results of Davenport–Heilbronn [13] and the first and second authors [9], respectively, answer the question in the related setting in which we consider $\text{GL}_2(\mathbb{Z})$-orbits on binary $n$-ic forms. However, for $n \geq 5$, it has not previously been known whether this density exists or even whether the lower density is positive. In this paper, we prove:

**Theorem 1** Let $n \geq 2$ be an integer. When integral binary $n$-ic forms $f(x,y) = a_0 x^n + a_1 x^{n-1} y + \cdots + a_n y^n$ are ordered by $H(f) := \max\{|a_0|, \ldots, |a_n|\}$, the density of forms having squarefree discriminant exists and is equal to

$$
\frac{1}{7} \prod_{p>2} \left( 1 - \frac{1}{p} \right) \left( 1 + \frac{1}{p} - \frac{1}{p^2} \right) \approx 38.97\% \quad \text{if } n = 2;
$$

$$
\frac{3}{7} \prod_{p>2} \left( 1 - \frac{1}{p} \right)^2 \left( 1 + \frac{1}{p^2} \right) \approx 24.64\% \quad \text{if } n = 3;
$$

$$
\frac{3}{7} \prod_{p>2} \left( 1 - \frac{1}{p} \right)^2 \left( 1 + \frac{2}{p} - \frac{2}{p^2} + \frac{1}{p^3} \right) \approx 21.18\% \quad \text{if } n = 4;
$$

$$
\frac{3}{7} \prod_{p>2} \left( 1 - \frac{1}{p} \right)^2 \left( 1 + \frac{1}{p} \right) \left( 1 + \frac{1}{p} - \frac{1}{p^2} \right) \approx 20.83\% \quad \text{if } n \geq 5.
$$
To any nonzero integral binary $n$-ic form $f(x, y) = a_0x^n + \cdots + a_ny^n$, we may naturally attach a rank-$n$ ring $R_f$ (see Birch–Merriman [11], Nakagawa [19], and Wood [30]), defined as follows when $a_0 \neq 0$. Let $\theta$ denote the image of $x$ in $K_f := \mathbb{Q}[x]/(f(x, 1))$. Let $R_f$ be the free rank-$n \mathbb{Z}$-submodule of $K_f$ generated by $1, a_0\theta, a_0\theta^2 + a_1\theta, \ldots, a_0\theta^{n-1} + \cdots + a_{n-1}\theta$. Then $R_f$ is in fact closed under multiplication and forms a ring whose discriminant is equal to the discriminant of $f(x)$. Our next result determines the density of irreducible integral binary forms $f$ for which $R_f$ is the maximal order in its field of fractions.

**Theorem 2** Let $n \geq 2$ be an integer. When irreducible integral binary $n$-ic forms $f(x, y) = a_0x^n + a_1x^{n-1}y + \cdots + a_ny^n$ are ordered by $H(f) := \max\{|a_0|, \ldots, |a_n|\}$, the density of forms $f$ such that $R_f$ is the ring of integers in $K_f$ exists and is equal to

\[
\prod_p \left(1 - \frac{1}{p^2} - \frac{1}{p^3} + \frac{1}{p^4}\right) \approx 53.59\% \quad \text{if } n = 2;
\]

\[
\zeta(2)^{-1}\zeta(3)^{-1} \approx 50.57\% \quad \text{if } n \geq 3.
\]

In particular, Theorem 2 yields the first unconditional Bertini theorem for arithmetic schemes of dimension $\geq 2$ as conjectured by Poonen [23, §5]. Indeed, for a quasiprojective subscheme $X$ of $\mathbb{P}^n_\mathbb{Z}$ that is regular of dimension $m$, Poonen conjectured that the density of hyperplane sections of $X$ that are regular of dimension $m - 1$ should equal $\zeta_X(m + 1)$, where $\zeta_X$ denotes the zeta function of $X$. Since the subscheme of $\mathbb{P}^1_\mathbb{Z}$ cut out by an integral binary $n$-ic form $f$ is regular if and only if $R_f$ is maximal, and the zeta function of $\mathbb{P}^1_\mathbb{Z}$ is given by $\zeta_{\mathbb{P}^1_\mathbb{Z}}(s) = \zeta(s)(s - 1)$, we have $\zeta_{\mathbb{P}^1_\mathbb{Z}}((\dim(\mathbb{P}^1_\mathbb{Z}) + 1)^{-1} = \zeta(2)^{-1}\zeta(3)^{-1}$. Therefore, Theorem 2 yields an unconditional proof of [23, Theorem 5.1] for the case $X = \mathbb{P}^1_\mathbb{Z}$ with the usual “box ordering” on the forms defining the hyperplane sections. In fact, we prove the stronger result that for every fixed $n \geq 3$, the density of regular binary $n$-ic forms is $\zeta(2)^{-1}\zeta(3)^{-1}$, while arithmetic Bertini only claims this in the limit as $n \to \infty$.

As a further application of our methods, we obtain the following theorem:

**Theorem 3** For each $n \geq 3$, the number of isomorphism classes of number fields of degree $n$ with associated Galois group $S_n$ and absolute discriminant less than $X$ is $\gg X^{1/2+1/(n-1)}$.

Our lower bound in Theorem 3 on the number of degree-$n$ $S_n$-number fields of absolute discriminant less than $X$ improves the previous best-known lower bound of $X^{1/2+1/n}$ obtained in [10]. We note that the number fields constructed in Theorem 3 can all be taken to have squarefree discriminant.

Our results also correct an error in, and thus resurrect, all the results of Nakagawa [19] and [21] that had been subsequently retracted in [20] and [22]. Specifically, the retracted theorems [19, Theorems 3–4] and [21, Theorem 2] regarding binary forms and $A_n$-extensions of quadratic fields can now be taken to be true. In particular, we obtain:

**Theorem 4** For $n \geq 3$, the total number of unramified $A_n$-extensions of real (resp., imaginary) quadratic fields $F$, across all such $F$ such that $|\text{Disc}(F)| < X$, is $\gg X^{(n+1)/(2n-2)}$.

Theorem 4 yields the best-known lower bounds on the number of unramified $A_n$-extensions of quadratic fields when $n > 5$. For improved bounds in the cases $n \leq 5$, see [3, Theorem 1.4]. For the best-known bounds on the number of quadratic fields of bounded discriminant admitting an unramified $A_n$-extension, see Kedlaya [18, Corollary 1.4]. Other related works include Uchida [29], Yamamoto [32], and Yamamura [33].
The main technical ingredient required to prove all the above results is a "tail estimate" which shows that not too many discriminants of integral binary $n$-ic forms $f$ are divisible by $p^2$ when $p$ is large relative to the discriminant of $f$ (here, large means larger than $H(f)$, say). It is these tail estimates that were missing in Nakagawa’s work. For a prime $p$, and an integral binary $n$-ic form $f$ such that $p^2 \mid \Delta(f)$, we say that $p^2$ strongly divides $\Delta(f)$ if $p^2 \mid \Delta(f + pg)$ for every integral binary $n$-ic form $g$; otherwise, we say $p^2$ weakly divides $\Delta(f)$. For any squarefree integer $m > 0$, let $\mathcal{W}_m^{(1)}$ (resp., $\mathcal{W}_m^{(2)}$) denote the set of integral binary $n$-ic forms whose discriminants are strongly divisible (resp., weakly divisible) by $p^2$ for every prime factor $p$ of $m$.

We prove the following tail estimates:

**Theorem 5** For an integer $n \geq 3$, a positive real number $M$ and any $\epsilon > 0$, we have:

(a) \[
\# \bigcup_{m > M \text{ squarefree}} \{ f \in \mathcal{W}_m^{(1)} : H(f) < X \} = O\left( \frac{X^{n+1+\epsilon}}{M} + X^n \right);
\]

(b) \[
\# \bigcup_{m > M \text{ squarefree}} \{ f \in \mathcal{W}_m^{(2)} : H(f) < X \} = O\left( \frac{X^{n+1+\epsilon}}{M} + X^{n+1-1/(5n)+\epsilon} \right), \text{ if } 2 \mid n;
\]

(c) \[
\# \bigcup_{m > M \text{ squarefree}} \{ f \in \mathcal{W}_m^{(2)} : H(f) < X \} = O\left( \frac{X^{n+1+1/(88n^5)}}{\sqrt{M}} + X^{n+1-1/(88n^6)} \right), \text{ if } 2 \mid n.
\]

The estimate in the strongly divisible Case (a) of Theorem 5 follows from geometric techniques, namely, the quantitative version of the Ekedahl geometric sieve as developed by the first author [3]. The estimates in the weakly divisible Cases (b) and (c) of Theorem 5 are considerably more difficult (particularly (c)), and we describe their proofs in the next section. Our tail estimate in fact allows us to prove Theorems 1 and 2 with power-saving error terms:

**Theorem 6** Let $V_n = \text{Sym}^n(2)$ denote the space of binary $n$-ic forms. Define $\eta_n = 1/(5n)$ when $n$ is odd and $1/(88n^6)$ when $n$ is even. Then

\[
\# \{ f \in V_n(\mathbb{Z}) : H(f) < X \text{ and } \Delta(f) \text{ squarefree} \} = \alpha_n \cdot (2X)^{n+1} + O(\frac{X^{n+1-\eta_n+\epsilon}}{X^n});
\]

\[
\# \{ f \in V_n(\mathbb{Z}) : H(f) < X \text{ and } R_f \text{ maximal} \} = \beta_n \cdot (2X)^{n+1} + O(\frac{X^{n+1-m+\epsilon}}{X^n}).
\]

These power saving bounds have applications towards level-of-distribution questions when counting integral binary $n$-ic forms $f$ of bounded height with $\Delta(f)$ squarefree (resp., $R_f$ maximal) satisfying splitting conditions at finitely many primes. Such level-of-distribution results in turn have applications towards a host of problems in analytic number theory, such as studying statistics of Artin $L$-functions attached to binary $n$-ic forms and proving lower bounds on the number of degree-$n$ number fields which are ramified only at a bounded number of primes, among many others.

We remark that our methods imply that the analogues of all of the above results also hold when local conditions are imposed at finitely many places (including at infinity); the orders of magnitudes in these theorems remain the same, provided that no local conditions are imposed that force the sets being counted in Theorems 1 and 2 to be empty.

Finally, the methods introduced in [10] and in the current article have applications beyond just squarefree values of polynomial discriminants. They have been recently adapted in [8] to determine the density of squarefree discriminants of elliptic curves over $\mathbb{Q}$ having two marked rational points. Other applications include determining the density of conductors in some families of elliptic curves [27] and the density of squarefree values taken by $a^4 + b^5$ ([24]).
2 Outline of proof

As mentioned in the introduction, the uniformity estimate in Theorem 5 is the key to deducing Theorems 1, 2, and 6 via a squarefree sieve. Case (a) of Theorem 5 follows directly from the results in [3]. Case (b), which pertains to odd degrees \( n \), can be proven using methods similar to those developed in our previous work [10]. However, these methods fail to work for Case (c), which pertains to even degrees \( n \), and a number of new ideas are required to handle this case. It is the proof of this case to which the bulk of our paper is devoted; it requires, in particular, the introduction of a new technique in the geometry of numbers that we believe may also be useful in other contexts.

In this section, we give a detailed outline of the proof of Case (b) pertaining to odd \( n \). We then explain why this strategy breaks down (quite spectacularly!) when \( n \) is even, and finally we describe the new techniques required to complete the proof of Theorem 5(c).

Sketch of the proof of the tail estimate for odd \( n \)

Our proof of Theorem 5(b) makes use of the representation of \( G = \text{SL}_n \) on the space \( W = 2 \otimes \text{Sym}_2(n) \) of pairs \( (A, B) \) of symmetric \( n \times n \) matrices, studied in detail in [31, 4, 6, 7]. The group \( G \) acts on \( W \) via \( \gamma \cdot (A, B) = (\gamma A \gamma^t, \gamma B \gamma^t) \) for \( \gamma \in G \) and \( (A, B) \in W \). We define the invariant binary form of an element \( (A, B) \in W \) by

\[
f_{A,B}(x, y) = (-1)^{(n-1)/2} \det(Ax - By).
\]

Then \( f_{A,B} \) is a binary \( n \)-ic form satisfying \( f_{\gamma(A,B)} = f_{A,B} \). Moreover, the ring of polynomial invariants for the action of \( G \) on \( W \) is freely generated by the coefficients of the invariant binary form (reference? it is in some classification of coregular spaces). Define the discriminant \( \Delta(A,B) \) and height \( H(A,B) \) of an element \( (A, B) \in W \) by \( \Delta(A,B) = \Delta(f_{A,B}) \) and \( H(A,B) = H(f_{A,B}) \).

The first step of our proof is the construction, for every squarefree integer \( m > 0 \), of a map

\[
\sigma_m : W_m^{(2)} \to W(\mathbb{Z}),
\]

such that \( f_{\sigma_m(f)}(x, y) = f(x, y) \) for every \( f \in W_m^{(2)} \). In our construction, the image of \( \sigma_m \) in fact lies in \( W_0(\mathbb{Z}) \), where \( W_0 \) is the subspace of \( W \) consisting of pairs of matrices whose top left \( g \times g \) blocks are zero, where \( n = 2g + 1 \). The action of the group \( G \) does not preserve \( W_0 \), and we take \( G_0 \) to be the maximal parabolic subgroup of \( G \) that does preserve \( W_0 \). When the discriminant polynomial \( \Delta \in \mathbb{Z}[W] \) is restricted to \( W_0 \), it is no longer irreducible but rather is divisible by the square of a polynomial \( Q \in \mathbb{Z}[W_0] \). This polynomial \( Q \) is a relative invariant for the action of \( G_0 \) on \( W_0 \). Its significance is that, by construction of \( \sigma_m \), every element in the image of \( \sigma_m \) has \( Q \)-invariant equal to \( m \). To prove Part (b) of Theorem 5, it therefore suffices to estimate the number of \( G_0(\mathbb{Z}) \)-orbits on \( W_0(\mathbb{Z}) \) having height less than \( X \) and \( Q \)-invariant greater than \( M \).

Bounding the number of these orbits is complicated by the fact that \( G_0 \) is not reductive. We are rescued by using the full action of \( G(\mathbb{Z}) \) on \( W(\mathbb{Z}) \). This necessitates expanding the definition of the \( Q \)-invariant from \( W_0(\mathbb{Z}) \) to all “distinguished” elements of \( W(\mathbb{Z}) \). An element \( (A, B) \in W(\mathbb{Z}) \) is distinguished if \( A \) and \( B \) have a common isotropic \( g \)-dimensional subspace defined over \( \mathbb{Q} \). Thus every element in \( W_0(\mathbb{Z}) \) (and thus every element in the image of \( \sigma_m \)) is distinguished. The \( Q \)-invariant, though defined initially on \( W_0 \), can be extended as a function on the set of all triples \( (A, B, \Lambda) \), where \( (A, B) \in W(\mathbb{Z}) \) is distinguished, and \( \Lambda \) is a common isotropic subspace of \( A \) and \( B \). For all but a negligible number of distinguished elements \( (A, B) \in W(\mathbb{Z}) \), \( A \) and \( B \) have exactly one common isotropic subspace \( \Lambda \) defined over \( \mathbb{Q} \). Thus we may define a \( G(\mathbb{Z}) \)-invariant function \( Q \) on the set of distinguished pairs \( (A, B) \in W(\mathbb{Z}) \) outside a negligible number of them. It then suffices
to bound the number of $G(\mathbb{Z})$-orbits on distinguished elements in $W(\mathbb{Z})$ having bounded height and large $Q$-invariant.

To obtain such a bound, we construct fundamental domains for the action of $G(\mathbb{Z})$ on elements in $W(\mathbb{R})$ with height less than $X$. Such a fundamental domain has a natural partition into three parts that we term the main body, the shallow cusp, and the deep cusp. We have little control over the $Q$-invariants of elements in the main body and the shallow cusp. However, it is known [17, Proposition 4.3] that there are a negligible number of integral elements in the shallow cusp. Meanwhile, distinguished elements occur rarely in the main body, a fact we prove via the Selberg sieve.

Finally, the deep cusp lies in $W_0$, where an upper bound for the $Q$-invariant can be obtained. Imposing the condition that this upper bound is greater than $M$, and counting the number of such points in the deep cusp using the averaging method of [2], gives the desired saving for the number of elements in the deep cusp having $Q$-invariant larger than $M$. Combining the estimates for the main body, the shallow cusp, and the deep cusp yields Part (b) of Theorem 5.

**Sketch of the proof of the tail estimate for even $n$**

With $W$ again denoting the space of pairs of symmetric $n \times n$ matrices, we may attempt to proceed in the same manner as in the case of odd $n$, by constructing a map

$$\sigma_m : W_m^{(2)} \to W(\mathbb{Z})$$

such that $f_{\sigma_m(f)}(x, y) = f(x, y)$ for every $f \in W_m^{(2)}$. However, such a map does not exist in the case that $n$ is even! Indeed, there exist integral binary $n$-ic forms $f(x, y)$ that cannot be expressed as $\det(Ax - By)$—even up to sign—for any integral $n \times n$ symmetric matrices $A$ and $B$. This phenomenon was extensively studied in [4, 6, 7]. It is in this sense that the strategy to prove Theorem 5(b) for odd $n$ fails spectacularly for even $n$—and at the very first step.

We address this issue by replacing $f(x, y) \in W_m^{(2)}$ by $xf(x, y)$, which is a reducible binary $(n+1)$-ic form whose discriminant, at least generically, remains weakly divisible by $m^2$. For these forms $xf(x, y)$, we can use the lift $\sigma_m$ constructed in the odd case. However, since $xf(x, y)$ has vanishing $y^{n+1}$ term, the image of $\sigma_m$ lies within the set of pairs $(A, B)$ where $B$ is singular.

The singularity of $B$ introduces additional difficulties with respect to both the algebraic and the analytic aspects of the proof. On the algebraic side, the main new problem is that distinguished elements $(A, B)$ with $B$ singular have at least two values for the $Q$-invariant, since they share at least two different common isotropic $(q + 1)$-dimensional subspaces, where $n = 2q + 2$. So it is no longer well-defined to impose the condition that $Q$ is large. Imposing the condition that the maximum value of $Q$ is large does not yield sufficient savings to prove an analogue of Theorem 5(b).

We thus instead construct a new invariant, termed $q$, such that for all but a negligible number of elements $(A, B)$ in the image of our map $\sigma_m$, the invariant $q$ is the minimum value taken by $Q$, and it satisfies $q(\sigma_m(xf(x, y))) = \pm m$.

As in the odd degree case, we once again construct fundamental domains $F_X$ for the action of $G(\mathbb{Z})$ on $W(\mathbb{R})$ with height less than $X$, and partition such a domain into three parts: the main body, the shallow cusp, and the deep cusp. However, we must now only count integer elements $(A, B)$ where $B$ is singular. The beautiful work of Eskin and Katznelson [15] provides asymptotics for the number of singular symmetric matrices in homogenously expanding domains. But this work is not directly applicable to our case since we need to estimate the number of singular symmetric matrices $B$ in skewed domains. To achieve this, we provide a simplification of the proof of the upper bounds in [15], at the cost of some extra log factors, which gives us a flexible method by which to obtain upper bounds on the number of singular symmetric matrices in arbitrarily skewed domains.
Accounting for the singularity of the $B$'s introduces complications in each region of the fundamental domain. In the main body, the lack of an exact count with a power-saving error term means we cannot directly apply a Selberg sieve to bound the number of distinguished elements. Instead, we fiber over the singular matrices $B$ and apply the Selberg sieve to bound the number of possible $A$'s. This requires us to prove new density estimates on the number of distinguished elements $(A,B)$ over $\mathbb{F}_p$, when $B$ is fixed.

Furthermore, unlike in the odd degree case, we no longer have an automatic power-saving on the number of pairs $(A,B) \in W(\mathbb{Z})$ lying in the shallow cusp of the fundamental domain and where $B$ is singular. As we go closer to the deep cusp, there are regions in which imposing the condition that $B$ is singular yields no saving whatsoever. To obtain the required bounds, we isolate this region of the shallow cusp and prove that integral elements $(A,B)$ in them either satisfy $\Delta(A,B) = 0$ or $|q(A,B)|$ is small.

Finally, for the deep cusp of $\mathcal{F}$, we once again use the condition that the $q$-invariant is large to obtain a power saving. Unlike the situation with the $Q$-invariant in the odd-degree case, the invariant $q$ in the even degree case behaves more wildly and is much harder to control. This is because $q$ is not a polynomial in the coefficients of $W_0$ but rather is a minimum of the different possible values of $Q$. In fact, there are regions within the deep cusp where the $q$-invariant of elements $(A,B)$ are not small. However, we show that these regions correspond to an archimedean condition on the invariant binary form $f$ of $(A,B)$, namely, that the discriminant of $f$ is much smaller than is typical for the height bound on $f$. Separately bounding the number of such binary forms yields the desired result.

Organization of the paper

This paper is organized as follows. We begin in §3 by recalling the arithmetic invariant theory for the representations $W_n := 2 \otimes \text{Sym}_2(n)$ of $\text{SL}_n$ and $2 \otimes g \otimes (g + 1)$ of $\text{SL}_2 \times \text{SL}_g \times \text{SL}_{g+1}$. In particular, we define the fundamental invariants $Q$ and $q$. We then construct our maps from $W_m^{(2)}$ into $W_n(\mathbb{Z})$ when $n$ is odd and into $W_{n+1}(\mathbb{Z})$ when $n$ is even.

The analytic parts of the paper are carried out in §4–6. In §4, we prove the tail estimates of Theorem 5 for odd degrees $n$ using geometry-of-numbers techniques. In §5, we carry out the necessary groundwork to count the number of singular symmetric matrices that lie in skewed domains. Using these results, we prove the tail estimates for even degrees $n$ in §6, completing the proof of Theorem 5. In §7, we deduce the main results, Theorems 1–4, from the tail estimates using a squarefree sieve, although the exact constants occurring in Theorems 1 and 2 remain conditional upon certain local density computations. Finally, in the Appendix, we compute the local densities of integral binary $n$-ic forms whose discriminants are indivisible by $p^2$ (resp., whose associated rings are maximal at $p$), thereby completing the proofs of Theorems 1 and 2.

3 Invariant theory on spaces associated to binary $n$-ic forms

Fix a positive integer $n$ and consider the space $V_n = \text{Sym}^n(2)$ of binary $n$-ic forms of degree $n$. The group $\text{SL}_2$ acts on $V_n$ via linear change of variables: we have $\gamma \cdot f(x,y) := f((x,y) \cdot \gamma)$ for $\gamma \in \text{SL}_2$ and $f \in V_n$.

Let $W_n = 2 \otimes \text{Sym}_2(n)$ denote the space of pairs of $n \times n$ symmetric matrices $(A,B)$. The group $\text{SL}_2 \times \text{SL}_n$ acts on $(A,B)$ via

$$(\gamma_2, \gamma_n) \cdot (A,B) = (\gamma_n A \gamma_2^t, \gamma_n B \gamma_2^t) \cdot \gamma_2^t.$$
There is a natural map \( W_n \rightarrow V_n \) given by

\[
(A, B) \mapsto f_{A,B} := (-1)^{n(n-1)/2} \det(Ax - By),
\]

sending an element of \( W_n \) to its invariant binary \( n \)-ic form. The ring of \( SL_n(\mathbb{C}) \)-invariant polynomials on \( W_n(\mathbb{C}) \) is freely generated by the coefficients of the invariant binary \( n \)-ic form.

### 3.1 Arithmetic invariant theory for the representation \( 2 \otimes \text{Sym}_2(n) \) of \( SL_n \)

First, let \( n = 2g + 1 \) be an odd integer with \( g \geq 1 \). We recall some of the arithmetic invariant theory of the representation \( W := W_n \) of \( SL_n \) and its map (1) to \( V := V_n \); see [6] for more details.

Let \( k \) be a field of characteristic not 2. For a binary \( n \)-ic form \( f(x, y) = a_0x^n + \cdots + a_ny^n \in V(k) \) with \( \Delta(f) \neq 0 \) and \( a_0 \neq 0 \), let \( C_f \) denote the smooth hyperelliptic curve \( z^2 = f(x, y)y \) of genus \( g \) viewed as a curve in the weighted projective space \( \mathbb{P}(1, 1, g+1) \). Let \( J_f \) denote the Jacobian of \( C_f \).

Then the stabilizer of an element \((A, B) \in W(k) \) with invariant binary form \( f(x, y) \) is isomorphic to \( J_f[2](k) \). The set of \( SL_n(k) \)-orbits on \( W(k) \) with invariant binary form \( f(x, y) \) maps injectively into \( H^1(k, J_f[2]) \). An element \((A, B) \) (or an \( SL_n(k) \)-orbit) is distinguished if \( \Delta(A, B) \neq 0 \) and there exists a \( g \)-dimensional subspace defined over \( k \) that is isotropic with respect to both \( A \) and \( B \). If \((A, B) \) is distinguished, then its \( SL_n(k) \)-orbit corresponds to the identity element of \( H^1(k, J_f[2]) \), and the set of these \( g \)-dimensional subspaces is in bijection with \( J_f[2](k) \).

Let \( W_0 \subset W \) be the subspace of pairs of matrices whose top left \( g \times g \) blocks are zero. Then elements \((A, B) \) in \( W_0(k) \) with nonzero discriminant are all distinguished since the \( g \)-dimensional subspace \( Y_g \) spanned by the first \( g \) basis vectors is isotropic with respect to both \( A \) and \( B \). Moreover, every distinguished element of \( W(k) \) is \( SL_n(k) \)-equivalent to some element in \( W_0(k) \) since \( SL_n(k) \) acts transitively on the set of \( g \)-dimensional subspaces of \( \mathbb{P}^{n-1}(k) \). Let \( G_0 \) be the maximal parabolic subgroup of \( SL_n \) consisting of elements \( \gamma \) that preserve \( Y_g \). Elements of \( W_0 \) have block matrix form

\[
(A, B) = \left( \begin{array}{ccc} 0 & A^{top} \\ (A^{top})^t & A_1 \end{array} \right), \left( \begin{array}{ccc} 0 & B^{top} \\ (B^{top})^t & B_1 \end{array} \right),
\]

where \( A^{top}, B^{top} \) are \( g \times (g + 1) \) matrices and \( A_1, B_1 \) are \((g + 1) \times (g + 1)\)-symmetric matrices.

Meanwhile, elements of \( G_0 \) have the block matrix form

\[
\gamma = \left( \begin{array}{cc} \gamma_1 & 0 \\ n & \gamma_2 \end{array} \right) \in \left( \begin{array}{cc} \text{GL}_g & 0 \\ M_{(g+1) \times g} & \text{GL}_{g+1} \end{array} \right).
\]

An element \( \gamma \in G_0 \) acts on the top right \( g \times (g + 1) \) block of elements of \( W_0 \) by

\[
\gamma(A^{top}, B^{top}) = (\gamma_1 A^{top} \gamma_2^t, \gamma_1 B^{top} \gamma_2^t)
\]

where we use the superscript “top” to denote the top right \( g \times (g + 1) \) block of an \( n \times n \) symmetric matrix. The action of \( G_0 \) on \( W_0 \) restricts to an action on the space \( U_g := 2 \otimes g \otimes (g + 1) \) of pairs of \( g \times (g + 1) \)-matrices. Moreover, the unipotent radical \( M_{(g+1) \times g} \) of \( G_0 \) acts trivially on \( U_g \). We study the invariant theory for this action more closely in the next subsection.

We will also need some results in the case when \( n = 2g + 2 \) is even in Section 6 (specifically in the proof of Lemma 6.7). Let \( f(x, y) = a_0x^n + \cdots + a_ny^n \in V(k) \) with \( \Delta(f) \neq 0 \) and \( a_0 \neq 0 \). Let \( L = k[x]/(f(x)) \). Let \( V_f(k) \) denote the set of \((A, B) \in W_n(k) \) with \( f_{A,B} = f(x, y) \). Then \( V_f(k) \) is nonempty if and only if \( f \in k^{x^2}N_{L/k}(L^x) \). Note in particular that if \( f(x, y) \in V(\mathbb{R}) \) is negative definite, so that \( L = \mathbb{R}[x]/(f(x)) \simeq \mathbb{C}^{n/2} \) and \( a_0 < 0 \), then \( V_f(\mathbb{R}) \) is empty. On the other hand, if \( k \) is a finite field of characteristic not 2, then \( V_f(k) \) is always nonempty and the number of \( SL_n(k) \)-orbits equals the number of even degree factorizations of \( f(x, y) \) over \( k \).
### 3.2 The representation $2 \otimes g \otimes (g+1)$ of $\text{SL}_2 \times \text{SL}_g \times \text{SL}_{g+1}$ and the $Q$-invariant

In this section, we collect some algebraic facts about the representation $U_g = 2 \otimes g \otimes (g+1)$ of the group $H_g := \text{SL}_2 \times \text{SL}_g \times \text{SL}_{g+1}$. We start with the following proposition.

**Proposition 3.1** The representation $U_g$ of $\mathbb{G}_m \times H_g$ is prehomogeneous, i.e., the action of $\mathbb{G}_m \times H_g$ on $U_g$ has a single Zariski open orbit. Furthermore, the stabilizer in $H_g(\mathbb{C})$ of an element in the open orbit of $U_g(\mathbb{C})$ is isomorphic to $\text{SL}_2(\mathbb{C})$.

**Proof:** We prove this by induction on $g$. The assertion is clear for $g = 1$, where the representation is that of $\mathbb{G}_m \times \text{SL}_2 \times \text{SL}_2$ on $2 \times 2$ matrices; the single relative invariant in this case is the determinant, and the open orbit consists of nonsingular matrices. For higher $g$, we note that $U_g$ is a castling transform of $U_{g-1}$ in the sense of Sato and Kimura [25]. As a result, the orbits of $\mathbb{G}_m \times \text{SL}_2 \times \text{SL}_g \times \text{SL}_{g-1}$ on $2 \otimes g \otimes (g-1)$ are in natural one-to-one correspondence with the orbits of $\mathbb{G}_m \times \text{SL}_2 \times \text{SL}_g \times \text{SL}_{g+1}$ on $2 \otimes g \otimes (2g - (g-1)) = 2 \otimes g \otimes (g+1)$, and under this correspondence, the open orbit in $U_{g-1}$ maps to an open orbit in $U_g$ (cf. [25]). Thus all the representations $U_g$ for the action of $\mathbb{G}_m \times H_g$ are prehomogeneous.

Note that castling transforms preserve stabilizers over $\mathbb{C}$. Since the generic stabilizer for the action of $H_1(\mathbb{C})$ on $U_1(\mathbb{C})$ is clearly isomorphic to $\text{SL}_2(\mathbb{C})$, it follows that this remains the generic stabilizer for the action of $H_g(\mathbb{C})$ on $U_g(\mathbb{C})$ for all $g \geq 1$. □

Since castling transforms also preserve polynomial invariants and their irreducibility ([25, Proposition 18]), it follows that the ring of polynomial invariants for this action of $H_g$ on $U_g$ is generated by an irreducible polynomial. We now give an explicit description of this invariant.

Write an element in $U_g = 2 \times g \times (g+1)$ as a pair $(A^{\text{top}}, B^{\text{top}})$ of $g \times (g+1)$ matrices. For $1 \leq i \leq g+1$, let $A_i$ and $B_i$ denote the $g \times g$-matrices obtained from $A^{\text{top}}$ and $B^{\text{top}}$, respectively, by deleting the $i$th column. Define the binary $(g+1)$-ic form $f_i(x, y)$ to be $(-1)^{i+1} \det(A_i x - B_i y)$. Consider the $(g+1) \times (g+1)$ matrix $C$ whose $(i, j)$-entry is the $j$th coefficient of $f_i(x, y)$. Taking the determinant of $C$ yields a polynomial $Q = Q(A^{\text{top}}, B^{\text{top}})$ in the coordinates of $U_g$. The polynomial $Q$ is the hyperdeterminant of the $2 \times g \times (g+1)$ matrix $(A^{\text{top}}, B^{\text{top}})$ (cf. [16, Theorem 3.18]). As a consequence, it is irreducible and invariant under the action of $H_g$ on $U_g$ and thus generates the ring of polynomials for the action of $H_g$ on $U_g$.

Let $n = 2g + 1$ again be an odd integer. We return to the representation $W_0$ of $G_0$. Given an element $(A, B) \in W_0$, recall that we obtain an element $(A^{\text{top}}, B^{\text{top}}) \in U_g$ by taking the top right $g \times (g+1)$ blocks of $A$ and $B$. We define the $Q$-invariant of $(A, B) \in W_0$ as the $Q$-invariant of $(A^{\text{top}}, B^{\text{top}})$:

$$Q(A, B) := Q(A^{\text{top}}, B^{\text{top}}).$$

Then the $Q$-invariant is a relative invariant for $G_0$. More precisely, for any $\gamma \in G_0$ in the block matrix form (3), we have

$$Q(\gamma \cdot (A, B)) = \det(\gamma_1)^{g+1} \det(\gamma_2)^g Q(A, B) = \det(\gamma_1) Q(A, B),$$

since $\det(\gamma_1) \det(\gamma_2) = 1$. If $\gamma \in G_0(\mathbb{Z})$, then we have $\det(\gamma_1) = \det(\gamma_2) = \pm 1$. Hence the absolute value $|Q|$ of $Q$ is an invariant for the action of $G_0(\mathbb{Z})$ on $W_0(\mathbb{Z})$.

### 3.3 Divisibility properties of $\Delta$ when restricted to $W_0$

Let $n = 2g + 1$ be an odd integer. Write the coordinates on $W_0$ as $a_{ij}, b_{ij}$ with $i, j$ in the appropriate ranges. Let $R$ denote the ring of regular functions of $W_0$ over $\mathbb{Z}$, i.e., $R = \mathbb{Z}[W_0] = \mathbb{Z}[a_{ij}, b_{ij}]$. 

8
Consider the discriminant polynomial $\Delta \in R$ given by $\Delta(A, B) := \Delta(f_{A,B})$. In this section, we prove that $Q^2 | \Delta$ as polynomials in $R$, along with another useful divisibility result.

Let $Z$ be the closed subvariety of $W_0$ consisting of elements $(A, B)$ with $\Delta(A, B) = 0$, and let $Y \subset Z$ denote the closed subvariety of $W_0$ consisting of elements $(A, B)$ such that $f_{A,B}$ is either divisible by the cube of a binary form with degree $\geq 1$ or the square of a binary form with degree $\geq 2$. Both of these varieties $Y$ and $Z$ are defined over $\mathbb{Z}$ and are clearly $\text{SL}_2 \times G_0$-invariant.

Our first result states that the variety in $W_0$ cut out by $Q = 0$ does not lie in $Y$.

**Proposition 3.2** Let $(A, B) = ((a_{ij}), (b_{ij})) \in W_0(R)$ be the generic element. Then

$$(A, B) \mod Q \notin Y(R/(Q)).$$

**Proof:** Fix an odd prime $p$. Let $f(x, y)$ be an element of $V(\mathbb{Z})$, such that the reduction of $f(x, y)$ modulo $p$ factors as $x^2 h(x, y)$, where $h$ is irreducible. In particular, $f(x, y) \mod p$ is not divisible by either the cube of a binary form with degree $\geq 1$, or the square of a binary form with degree $\geq 2$. Let $(A_f, B_f) \in W_0(\mathbb{Z})$ be an element with invariant binary $n$-ic form equal to $f$ and $Q(A_f, B_f) = p$. Such an element $(A_f, B_f)$ is constructed in the next subsection.

Let $\pi : R \to \mathbb{Z}$ denote the specialization map assigning integer values to $a_{ij}, b_{ij}$ such that

$$\pi(A, B) = (A_f, B_f).$$

Then $\pi(Q) = p$ and so $\pi$ induces a map $R/(Q) \to \mathbb{F}_p$. Since $(A_f, B_f) \mod p \notin Y(\mathbb{F}_p)$, we see that $(A, B) \mod Q \notin Y(R/(Q)).$ $\square$

The next lemma, which follows from a direct computation, gives the $Q$-invariant for elements in $W_0$ having a specific form.

**Lemma 3.3** Let $k$ be a field and let $(A, B) \in W_0(k)$ be an element such that the top right $g \times (g+1)$ blocks of $(A, B)$ are of the following form:

$$\begin{pmatrix}
0 & 0 & \cdots & 0 & 0 & a_1 \\
0 & a_2 & * & & & \\
0 & a_3 & * & * & & \\
\vdots & \vdots & \vdots & \ddots & \ddots & \\
0 & a_g & \cdots & * & * & *
\end{pmatrix},
\begin{pmatrix}
0 & \cdots & 0 & 0 & b_1 & 0 \\
& b_2 & 0 & 0 & & \\
& b_3 & 0 & 0 & & \\
& \ddots & \ddots & \ddots & \ddots & \\
& b_g & \cdots & * & * & 0
\end{pmatrix}.$$  \hspace{1cm} (6)

Then

$$Q(A, B) = \pm a_1^g a_2^{g-1} \cdots a_g b_1 b_2^2 \cdots b_g^g.$$ 

Next, we have the following proposition that gives a normal form for elements $(A, B) \notin Y$ whose $Q$-invariant is 0.

**Proposition 3.4** Let $k$ be a field. Let $(A, B)$ be an element of $W_0(k) \setminus Y(k)$ such that $Q(A, B) = 0$. Then $(A, B)$ is $\text{SL}_2(k) \times G_0(k)$-equivalent to an element of the form $(A', B')$ where the top right $g \times (g+1)$ blocks of $A'$ and $B'$ are given by

$$\begin{pmatrix}
0 & 0 & \cdots & 0 & 0 & a_1 \\
0 & a_2 & * & & & \\
0 & a_3 & * & * & & \\
\vdots & \vdots & \vdots & \ddots & \ddots & \\
0 & a_g & \cdots & * & * & *
\end{pmatrix},
\begin{pmatrix}
0 & \cdots & 0 & 0 & 0 & 0 \\
& b_2 & 0 & 0 & & \\
& b_3 & 0 & 0 & & \\
& \ddots & \ddots & \ddots & \ddots & \\
& b_g & \cdots & * & * & 0
\end{pmatrix}.$$  \hspace{1cm} (7)

where $a_1, \ldots, a_g, b_2, \ldots, b_g \in k^*$. In the displayed matrices above, any empty entry is 0.
Proof: The action of $G_0(k)$ allows us to perform simultaneous row operations and simultaneous column operations on $(A^{\text{top}}, B^{\text{top}})$. As a first step, we perform column operations to ensure that the rightmost column of $B^{\text{top}}$ is 0. Next, recall that the $Q$-invariant of $(A, B)$ is the determinant of the $(g + 1) \times (g + 1)$ matrix $C$, whose rows come from the coefficients of the $g \times g$ minors of $A^{\text{top}}x - B^{\text{top}}y$. It follows that row operations on $(A^{\text{top}}, B^{\text{top}})$ leave $C$ unchanged, while adding $\alpha$ times the $i$-th columns of $A^{\text{top}}, B^{\text{top}}$ to the $j$-th column has the effect of adding $\alpha$ times the $j$-th row of $C$ to the $i$-th row of $C$ and leaving the rest unchanged. Since $\det(C) = Q(A^{\text{top}}, B^{\text{top}}) = 0$, it follows that by adding multiples of the last columns of $A^{\text{top}}, B^{\text{top}}$ to the other columns, we may assume that the last row of $C$ is 0. Denoting the $g \times g$ matrices obtained by removing the last columns of $A^{\text{top}}$ and $B^{\text{top}}$ by $M$ and $N$, respectively, we have $\det(Mx - Ny) = 0$.

We next claim that by performing simultaneous row and column operations on $(M, N)$, we may bring $M$ and $N$ in the form of the first $g$ columns of $A'^{\text{top}}$ and $B'^{\text{top}}$, respectively, for $(A'^{\text{top}}, B'^{\text{top}})$ as given in (7) with $b_i \neq 0$ for all $2 \leq i \leq g$. Since $\det(M) = 0$, after appropriate column operations we may assume that the first column of $M$ is 0. Now the first column of $N$ cannot be identically 0 for otherwise, the invariant binary form of $(A, B)$ has a factor of the form $h(x, y)^2$ with $\deg h = g$, contradicting $(A, B) \notin Y(k)$. By applying row operations, we may ensure that the bottom left entry of $N$ is $b_g \neq 0$ and the rest of the first column of $N$ is 0. We then use this nonzero coefficient $b_g$ to clear out the rest of the bottom row of $N$ (without changing $M$).

Let $M_1$ and $N_1$ denote the top right $(g-1) \times (g-1)$ block of $M$ and $N$. Then $\det(Mx - Ny) = (-1)^gb_0y\det(M_1x - N_1y)$. Hence $\det(M_1x - N_1y) = 0$ and the first column of $M_1$ can be made 0. As in the previous case, all the coefficients of the first column of $N_1$ can be made 0 except for the bottom left entry, which is $b_{g-1} \neq 0$. We then clear out the bottom row of $N_1$ as before. Proceeding in this way, we transform the first $g - 1$ columns of $M$ and $N$ to be in the required form. Since the $b_i$’s are nonzero for $2 \leq i \leq g$, and since $\det(Mx - Ny) = 0$, it follows that the top right coefficients of $M$ and $N$ are 0, completing the proof of the claim.

Note that this transformation of $M$ and $N$ did not change the last column of $B^{\text{top}}$, which remains 0. Thus to complete the proof of Proposition 3.4, it remains to show that $a_i \neq 0$ for $1 \leq i \leq g$. Since the first row and column of $B'$ are 0, we see that $x^2a_1^2 | f_{A', B'}$. Hence $a_1 \neq 0$. Suppose for contradiction that $i = 2, \ldots, g$ is the smallest index such that $a_i = 0$. Then we may clear out the $i$-th row of $A'$ using the second up to the $(i - 1)$-th rows of $A'$. That is, $(A', B')$ is $\text{SL}_g(k)$-equivalent to some $(A'', B'')$ where the only nonzero entries in the $i$-th row and the $i$-th column of $A''$ appear in the last entry. This allows us to factor out an extra factor of $y^2$ in $\det(A''x - B''y) = \pm f_{A, B}$, contradicting the assumption that $(A, B) \notin Y(k)$ since we already had $x^2 | f_{A, B}$. □

We are now ready to prove that $Q^2 \mid \Delta$:

**Theorem 3.5** We have $Q^2 \mid \Delta$ in $\mathbb{Z}[W_0]$.

**Proof:** Let $(A, B) \in W_0(R)$ be the generic element. We begin by proving that $(A, B) \in Z(R/(Q))$, or equivalently that $Q \mid \Delta$ in $R$. Let $(\bar{A}, \bar{B}) \in W_0(R/(Q))$ denote the reduction of $(A, B)$ mod $Q$ and let $F$ denote the field of fractions of $R/(Q)$. By Proposition 3.2, we know $(\bar{A}, \bar{B}) \notin Y(F)$. Since $Q(\bar{A}, \bar{B}) = 0$, by Proposition 3.4, there exists $\gamma \in \text{SL}_2(F) \times G_0(F)$ such that $\gamma(\bar{A}, \bar{B}) = (A', B')$ where $(A'^{\text{top}}, B'^{\text{top}})$ is of the form (7). The invariant binary form of $(A', B')$ has a factor of $x^2$, and so $(A', B') \in Z(F)$. Since $Z$ is $\text{SL}_2 \times G_0$-invariant, we see that $(\bar{A}, \bar{B}) \in Z(F)$.

Since $Q \mid \Delta$ in $R$, there exists an element $\delta \in R$ such that $\Delta = Q\delta$. Let $Z_1$ denote the closed subvariety of $W_0$ cut out by $\delta$. It now suffices to prove that $Q \mid \delta$ or, equivalently, that the generic element $(A, B)$ belongs to $Z_1(R/Q)$. We claim that for any field $k$, and every element
\((A, B) \in W_0(k)\) such that \((A^{\top}, B^{\top})\) has the form (7), we have \(\delta(A, B) = 0\). Indeed, let \((A, B)\) be such an element. Let \((A^{(\epsilon)}, B^{(\epsilon)}) \in W_0(k[\epsilon])\) be such that \(A^{(\epsilon)} = A\), the \((1, n-1)\)-entry and the \((n-1, 1)\)-entry of \(B^{(\epsilon)}\) equal \(\epsilon\), and the other coefficients of \(B^{(\epsilon)}\) are the same as those of \(B\). By Lemma 3.3, we have

\[
Q(A^{(\epsilon)}, B^{(\epsilon)}) = \pm \epsilon a_1^g a_2^{g-1} \cdots a_g b_2 \cdots b_g.
\]

Moreover, \(\epsilon^2\) divides the \(y^n\)-coefficient of \(f_{A^{(\epsilon)}, B^{(\epsilon)}}\) and \(\epsilon\) divides the \(xy^{n-1}\)-coefficient of \(f_{A^{(\epsilon)}, B^{(\epsilon)}}\). Hence \(\epsilon^2 \mid \Delta(A^{(\epsilon)}, B^{(\epsilon)})\), which implies (since \(\epsilon^2 \mid Q(A^{(\epsilon)}, B^{(\epsilon)})\)) that \(\epsilon \mid \delta(A^{(\epsilon)}, B^{(\epsilon)})\). Since \((A, B)\) is obtained from \((A^{(\epsilon)}, B^{(\epsilon)})\) by setting \(\epsilon = 0\), we have \(\delta(A, B) = 0\). We have proven that the generic element \((A, B) \in W_0(R)\) belongs to \(Z_1(R/(Q))\). Therefore, \(Q \mid \delta\). \(\square\)

We end this section with another divisibility result for \(\Delta\), which will be used in §6.

**Proposition 3.6** We have \(\det(A^{\top}(A^{\top})^t) \det(B^{\top}(B^{\top})^t) \mid \Delta\) as elements in \(Z[W_0]\).

**Proof:** It suffices to prove that \(\det(B^{\top}(B^{\top})^t)\) divides \(\Delta\) in \(Z[W_0]\). Suppose \(A = B \in W_0(\mathbb{C})\) with \(\det(B^{\top}(B^{\top})^t) = 0\). Then \(B^{\top}\) does not have full rank. Hence there exists some nonzero \(v \in \text{Span}_\mathbb{C}\{e_1, \ldots, e_g\}\) such that \(Bv = 0\). However, any such \(v\) is isotropic with respect to \(A\). As a result, \(\Delta(A, B) = 0\). Thus, by the Nullstellensatz, \(\det(B^{\top}(B^{\top})^t) \mid c\Delta^d\) in \(Z[W_0]\) for some nonzero integer \(c\) and positive integer \(d\).

Define \(P_g \in Z[M_{g \times (g+1)}]\) by \(P_g(M) = \det(MM^t)\). For the purpose of proving Proposition 3.6, it suffices to prove that \(P_g\) is squarefree in \(Z[M_{g \times (g+1)}]\). We proceed by induction on \(g\). Denote the \((i, j)\)-entry of any \(M \in M_{g \times (g+1)}\) by \(u_{ij}\). When \(g = 1\), we have \(P_1 = u_{11}^2 + u_{12}^2\), which is squarefree in \(Z[u_{11}, u_{12}]\). For general \(g \geq 2\), consider

\[
M = \begin{pmatrix}
u_{11} & \cdots & u_{1g-1} & u_{1g} & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots \\
u_{g-11} & \cdots & u_{g-1g-1} & u_{g-1g} & 0 \\
0 & \cdots & 0 & \alpha & \beta
\end{pmatrix}.
\]

Then

\[
\det(MM^t) = \beta^2 P_{g-1} + \alpha^2 D_{g-1}.
\]

where \(D_{g-1}\) is the determinant of the top left \((g - 1) \times (g - 1)\) block of \(M\). Any square factor of \(\det(MM^t)\) must be a common square factor of \(P_{g-1}\) and \(D_{g-1}^2\), which can only be \(\pm 1\) since \(P_{g-1}\) is squarefree by induction. We have shown that \(P_g\) is squarefree even after setting certain variables to 0. Therefore, \(P_g\) is squarefree in \(Z[M_{g \times (g+1)}]\). \(\square\)

### 3.4 Embedding \(W^{(2)}_{m,n}\) into \(W_n(\mathbb{Z})\), for \(n\) odd

Let \(n = 2g + 1\) be an odd integer, and set \(W = W_n\). For an odd squarefree integer \(m > 0\), let \(W_m^{(2)} = W_m^{(2)}(\mathbb{Z})\) denote the set of integer binary forms whose discriminants are weakly divisible by \(p^2\) for every prime factor \(p\) of \(m\). Fix an element \(f(x, y) \in W^{(2)}_m\). Then just as shown in [10, §3.2], there exists an \(\text{SL}_2(\mathbb{Z})\)-change of variable such that \(f((x, y)\gamma)\) has the form

\[
f((x, y)\gamma) = m^2 b_0 x^n + mb_1 x^{n-1} y + \cdots + b_n y^n
\]

for some integers \(b_0, \ldots, b_n\) and where \(m\) and \(b_0\) are coprime.
Consider the following pair of matrices:

\[
A = \begin{pmatrix}
1 & \cdots & \cdots & 1 \\
\vdots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
1 & \cdots & c_{n-1} & m \\
\end{pmatrix}, \quad B = \begin{pmatrix}
1 & \cdots & 1 \\
\vdots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & c_3 & c_1 & 1 \\
\end{pmatrix}.
\]

Here the dots on the antidiagonal of \(A\) are all 1. We claim that \(c_i, r\) can be chosen to be integers so that \((-1)^g \det(xA - yB) = f((x, y)\gamma)\). It is clear that \(c_0 = b_0\) and \(2mrc_0 + m^2c_1 = -mb_1\). Choose \(r \in \mathbb{Z}\) such that \(m \mid 2rc_0 + b_1\); this then determines \(c_1\). It is then not hard to check that the coefficient of \(x^{m-1}y^r\) in \((-1)^g \det(xA - yB)\) is of the form \((-1)^g c_1 + L(c_0, \ldots, c_{i-1})\) where \(L\) is a linear form with coefficients in \(\mathbb{Z}[r]\). The existence of integers \(c_2, \ldots, c_n\) now follows by induction.

Set \(\sigma_m(f) = \sigma_{m,n}(f)\) to be the element \((A_f, B_f)\) such that

\[\begin{pmatrix}
A_f \\
-B_f
\end{pmatrix} = \gamma^{-1}\begin{pmatrix}
A \\
-B
\end{pmatrix}.\]

Then \(f_{\sigma_m(f)} = f\). Next, we note that \((A, B)\) and thus \((A_f, B_f)\) are in \(W_0(\mathbb{Z})\) and from Lemma 3.3, we obtain that \(|Q|(A, B) = m\). Since \(Q\) is \(\text{SL}_2\)-invariant, we conclude that

\[|Q|(\sigma_m(f)) = m.\]

We have proven the following theorem.

**Theorem 3.7** Let \(m > 0\) be a squarefree integer. There exists a map \(\sigma_m : W_m^{(2)} \to W_0(\mathbb{Z})\) such that

\[f_{\sigma_m(f)} = f, \quad |Q|(\sigma_m(f)) = m\]

for every \(f \in W_m^{(2)}\).

We will later use the image of \(\sigma_1\) as a fundamental set for the action of \(\text{SL}_n(\mathbb{R})\) on the set of distinguished elements of \(W(\mathbb{R})\). We now extend the function \(|Q|\) to the set of distinguished elements of \(W(\mathbb{Z})\) having irreducible invariant binary form. Suppose that \((A, B)\) is a distinguished element of \(W(\mathbb{Z})\). Then there is a \(g\)-dimensional subspace \(X\) isotropic with respect to \(A\) and \(B\). Let \(\Lambda = X \cap \mathbb{Z}^n\) be the primitive lattice in \(X\). There exists an element \(\gamma\) in \(\text{SL}_n(\mathbb{Z})\), unique up to left multiplication by an element in \(G_0(\mathbb{Z})\), such that \(\Lambda = \gamma^t \cdot \text{Span}_\mathbb{Z}\{e_1, \ldots, e_n\}\), where \(e_1, \ldots, e_n\) is the standard basis of \(\mathbb{Z}^n\). Then \(\gamma \cdot (A, B) \in W_0(\mathbb{Z})\), and we can thus define the \(|Q|\)-invariant on the triple \((A, B, \Lambda)\) by

\[|Q|(A, B, \Lambda) := |Q|(\gamma \cdot (A, B)).\]

That is, we complete an integral basis of \(\Lambda\) to an integral basis of \(\mathbb{Z}^n\) with respect to which the pair \((A', B')\) of Gram matrices for the quadratic forms defined by \(A\) and \(B\) lies inside \(W_0(\mathbb{Z})\), and we define \(|Q|(A, B, \Lambda)\) to be \(|Q|(A', B')\).

We end with the following result that will be crucial in Section 4.
Proposition 3.8 Let $n = 2g + 1$ be an odd integer with $n \geq 3$. Let $m$ be an odd positive squarefree integer. Let $f(x, y) \in \mathcal{W}^{(2)}_m$ be an irreducible integral binary $n$-ic form. Let $(A, B)$ be any element in $\text{SL}_n(\mathbb{Z})\cdot \sigma_m(f)$. Then there is a unique primitive $g$-dimensional lattice $\Lambda$ that is isotropic with respect to both $A$ and $B$. Moreover, $|Q|(A, B) := |Q|(A, B, \Lambda) = m$. In particular, if $f(x, y) \in \mathcal{W}^{(2)}_m \cap \mathcal{W}^{(2)}_{m'}$ is irreducible where $m$ and $m'$ are distinct odd positive squarefree integers, then $\sigma_m(f(x, y))$ and $\sigma_{m'}(f(x, y))$ are not $\text{SL}_n(\mathbb{Z})$-equivalent.

Proof: Let $C_f$ denote the smooth hyperelliptic curve $z^2 = f(x, y)y$ of genus $g$ viewed as a curve in the weighted projective space $\mathbb{P}(1, 1, g + 1)$, and let $J_f$ denote its Jacobian. Since $(A, B)$ is $\text{SL}_n(\mathbb{Z})$-equivalent to $\sigma_m(f)$, it follows that $(A, B)$ is distinguished. Thus the set of common isotropic $g$-dimensional subspaces of $A$ and $B$ over $\mathbb{Q}$ is in bijection with $J_f[2](\mathbb{Q})$. Since $f$ is irreducible, we have $J_f[2](\mathbb{Q}) = 1$. Therefore, there is a unique primitive $g$-dimensional lattice $\Lambda$ which is isotropic with respect to both $A$ and $B$.

Let $\gamma \in \text{SL}_n(\mathbb{Z})$ be an element such that $\gamma(A, B) = \sigma_m(f) = (A_f, B_f) \in W_0(\mathbb{Z})$. Since we know that $\text{Span}_\mathbb{Z}\{e_1, \ldots, e_g\}$ is a primitive $g$-dimensional lattice isotropic with respect to $A_f$ and $B_f$, we see that $\gamma^t \cdot \text{Span}_\mathbb{Z}\{e_1, \ldots, e_g\}$ is a primitive $g$-dimensional lattice isotropic with respect to $A$ and $B$. By uniqueness, it follows that $\Lambda = \gamma^t \cdot \text{Span}_\mathbb{Z}\{e_1, \ldots, e_g\}$, and so by definition, $|Q|(A, B, \Lambda) = |Q|(\sigma_m(f)) = m$, where the final equality is Theorem 3.7. \(\square\)

3.5 Embedding $\mathcal{W}^{(2)}_{m,n}$ into $W_{n+1}(\mathbb{Z})$, for $n$ even

Suppose now that $n = 2g + 2$ is even with $g \geq 1$. For an odd squarefree integer $m > 0$, let $\mathcal{W}^{(2)}_{m,n}$ denote the set of integer binary forms having discriminant weakly divisible by $p^2$ for every prime factor $p$ of $m$. Let $\mathcal{W}^{(2),\text{gen}}_{m,n} \subset \mathcal{W}^{(2)}_{m,n}$ consist of those $f(x, y)$ with $f(0, 1)$ coprime to $m$. Since $\Delta(xf(x, y)) = \Delta(f(x, y))f(0, 1)^2$, we see that if $f(x, y) \in \mathcal{W}^{(2),\text{gen}}_{m,n}$, then $xf(x, y) \in \mathcal{W}^{(2)}_{m,n+1}$. We define $\sigma_{m,n} : \mathcal{W}^{(2),\text{gen}}_{m,n} \to W_{n+1}(\mathbb{Z})$ via $\sigma_{m,n}(f) = \sigma_{m,n+1}(xf)$. For the rest of this subsection, we drop subscripts and denote $\sigma_{m,n}$ by $\sigma_m$, $\mathcal{W}^{(2),\text{gen}}_{m,n}$ by $\mathcal{W}^{(2),\text{gen}}_m$, and $W_{n+1}$ by $W$.

We now define the finer $q$-invariant. Let $f \in \mathcal{W}^{(2),\text{gen}}_m$ and suppose $(A, B) = \sigma_m(f)$. Then $B$ is singular since $xf(x, y)$ has vanishing $y^n$-term. Moreover, since $\Delta(xf) \neq 0$, the kernel of $B$ has dimension exactly $1$ and is not isotropic with respect to $A$ (see Lemma 6.4). Fix an integral domain $D$. Let $W_1(D)$ be the subset of $W(D)$ consisting of pairs $(A, B)$ of symmetric $(n+1)\times(n+1)$ matrices satisfying the following conditions:

(a) The top left $(g+1) \times (g+1)$ block of $A$ is $0$.
(b) The top left $(g+2) \times (g+2)$ block of $B$ is $0$ (implying that $B$ is singular).
(c) The kernel of $B$ has dimension exactly $1$ (over the fraction field of $D$) and is not isotropic with respect to $A$.

Take any $(A, B) \in W_1(D)$. Conditions (a) and (c) imply that the first $g+1$ columns of $B$ are linearly independent over the fraction field of $D$. Let $B'$ denote the top right $(g+1) \times (g+1)$ block of $B$. Since $B$ is symmetric, we see that $B'$ is nonsingular. It is now easy to see from the definition of the $Q$-invariant that as polynomials in the coordinates of $W_1(D)$, we have

$$\det(B') | Q(A^{\text{top}}, B^{\text{top}}).$$

We define the quotient to be the $q$-invariant of $(A, B)$:

$$q(A, B) := Q(A^{\text{top}}, B^{\text{top}})/\det(B').$$

(9)
Let $G_1(D)$ denote the subgroup of $\text{SL}_{n+1}(D)$ preserving $W_1(D)$. Then elements of $G_1(D)$ have the following block matrix form

$$
\gamma = \begin{pmatrix} \gamma_1 & 0 & 0 \\ n_1 & \gamma_2 & 0 \\ n_2 & n_3 & \gamma_3 \end{pmatrix} \in \begin{pmatrix} \text{GL}_{g+1} & M_{1 \times (g+1)} & \text{GL}_1 \\ M_{(g+1) \times (g+1)} & M_{(g+1) \times 1} & \text{GL}_{g+1} \end{pmatrix}.
$$

(10)

It is easy to check that for any $(A, B) \in W_1(D)$,

$$
q(\gamma(A, B)) = \det(\gamma_1)(\det(\gamma_1)\det(\gamma_3))^{-1}q(A, B) = \det(\gamma_1)\gamma_2 q(A, B).
$$

(11)

We now consider the situation over $\mathbb{Z}$. Let $(A, B) \in W(\mathbb{Z})$ be a distinguished element having nonzero discriminant such that $B$ is singular. Let $X$ denote a common isotropic $(g+1)$-dimensional subspace of $A$ and $B$. We know that the kernel $\langle v \rangle$ of $B$ has trivial intersection with $X$. Denote the span of $X$ and $v$ by $X'$, which is a $(g+2)$-dimensional subspace containing $X$ that is isotropic with respect to $B$. Let $\Lambda = X \cap \mathbb{Z}^{n+1}$ and $\Lambda' = X' \cap \mathbb{Z}^{n+1}$ be the primitive lattices in $X$ and $X'$, respectively. There exists an element $\gamma$ in $\text{SL}_{n+1}(\mathbb{Z})$, unique up to left multiplication by an element in $G_1(\mathbb{Z})$, such that $\Lambda = \gamma' \cdot \text{Span}_\mathbb{Z}\{e_1, \ldots, e_{g+1}\}$ and $\Lambda' = \gamma' \cdot \text{Span}_\mathbb{Z}\{e_1, \ldots, e_{g+2}\}$. Then $\gamma(A, B) \in W_1(\mathbb{Z})$, and we can thus define the $|q|$-invariant for the quadruple $(A, B, \Lambda, \Lambda')$ by

$$
|q|(A, B, \Lambda, \Lambda') := |q(\gamma(A, B))|.
$$

In other words, we complete an integral basis $\{v_1, \ldots, v_{g+1}\}$ of $\Lambda$ to an integral basis $\{v_1, \ldots, v_{n+1}\}$ of $\mathbb{Z}^{n+1}$ such that $\{v_1, \ldots, v_{g+2}\}$ forms an integral basis of $\Lambda'$. When expressed in this basis, the pair $(A', B')$ of Gram matrices for the quadratic forms defined by $A$ and $B$ lies in $W_1(\mathbb{Z})$ and we define $|q|(A, B, \Lambda, \Lambda') := |q(A', B')|$.

Finally, we compute the $|Q|$- and $|q|$-invariants of $\sigma_m(f(x, y))$, where $f(x, y) \in W_m^{(2), \text{gen}}$ is irreducible.

**Proposition 3.9** Let $n = 2g + 2$ with $g \geq 1$. Let $m$ be an odd positive squarefree integer. Let $f(x, y) \in W_m^{(2), \text{gen}}$ be irreducible. Let $(A, B)$ be any element in $\text{SL}_{n+1}(\mathbb{Z}) \cdot \sigma_m(f(x, y))$. Let $\Lambda$ be a $(g+1)$-dimensional primitive lattice contained in a $(g+2)$-dimensional primitive lattice $\Lambda'$ such that $\Lambda$ is isotropic with respect to $A$ and $\Lambda'$ is isotropic with respect to $B$. Then $|Q|(A, B, \Lambda)$ is either $m$ or $|f(0, 1)|m$, and $|q|(A, B) := |q|(A, B, \Lambda, \Lambda') = m$, independent of $(\Lambda, \Lambda')$. In particular, if $f(x, y) \in W_m^{(2), \text{gen}} \cap W_m^{(2), \text{gen}}$ is irreducible where $m$ and $m'$ are distinct odd positive squarefree integers, then $\sigma_m(f(x, y))$ and $\sigma_{m'}(f(x, y))$ are not $\text{SL}_{n+1}(\mathbb{Z})$-equivalent.

**Proof:** The size of $J_f[2](\mathbb{Q})$ is 2 since $xf(x, y)$ has a unique even degree factor (namely, $f(x, y)$) over $\mathbb{Q}$. Therefore, the pair $(A, B)$ has two $(g+1)$-dimensional common isotropic subspaces $X_1$ and $X_2$ over $\mathbb{Q}$. Let $\Lambda_1$ and $\Lambda_2$ denote the corresponding primitive lattices contained in $X_1$ and $X_2$. The unique $(g+2)$-dimensional subspace $X'_1$ (resp., $X'_2$) isotropic with respect to $B$ and containing $X_1$ (resp., $X_2$) is the span of $X_1$ (resp., $X_2$) with the kernel of $B$. Let $\Lambda'_1$ and $\Lambda'_2$ denote the primitive lattices contained in $X'_1$ and $X'_2$. We compute the $|Q|$- and $|q|$-invariants associated to these lattices.

We may assume that $(A, B) = \sigma_m(f)$ since the action of $\text{SL}_{n+1}(\mathbb{Z})$ does not change the $|Q|$- or $|q|$ invariants. Since $|Q|$ is $\text{SL}_2$-invariant, and $|q|$ remains unchanged when we add a multiple of $B$ to $A$, we may also assume that

$$
xf(x, y) = m^2b_0x^{n+1} + mb_1x^ny + \cdots + b_nxy^n
$$

14
and so

\[
A = \begin{pmatrix}
1 & m & \\
m & c_0 & m \\
0 & c_1 & c_2 \\
& & \ddots & \ddots & c_n \\
& & & 1 & \\
\end{pmatrix}, \quad B = \begin{pmatrix}
1 & r & \\
r & c_1 & r \\
0 & c_2 & c_3 \\
& & \ddots & \ddots & \ddots \\
\end{pmatrix}
\]

Comparing the \(y^{n+1}\)- and the \(xy^n\)-coefficients, we have \(c_{n+1} = 0\) and \(c_n = b_n\). Let \(\langle , \rangle_A\) and \(\langle , \rangle_B\) denote the quadratic forms corresponding to \(A\) and \(B\). Let \(e_1, \ldots, e_{n+1}\) be the standard basis on \(\mathbb{Z}^{n+1}\). Since \(c_{n+1} = 0\), the vector \(e_{n+1}\) spans the kernel of \(B\). We may take the subspace spanned by \(e_1, \ldots, e_g\) as \(X_1\). Then by construction \(|Q|(A, B, \Lambda_1) = m\). When expressed in terms of the ordered integral basis \(\{e_1, \ldots, e_g, e_{g+1}, \ldots, e_{n-1}\}\), the top right \((g+1) \times (g+1)\) block of \(B\) has 1’s on the antidiagonal and 0’s above the antidiagonal, and so has determinant \(\pm 1\). Hence \(|q|(A, B, \Lambda_1, \Lambda_1') = m\).

The second common isotropic \((g+1)\)-dimensional subspace \(X_2\) is the reflection of \(X_1\) in the hyperplane perpendicular to \(e_{n+1}\) with respect to \(\langle , \rangle_A\). That is,

\[
X_2 = \text{Span}_Q \left\{ e_1 - \frac{2(e_1, e_{n+1})_A}{(e_{n+1}, e_{n+1})_A} e_{n+1}, \ldots, e_{g+1} - \frac{2(e_{g+1}, e_{n+1})_A}{(e_{n+1}, e_{n+1})_A} e_{n+1} \right\} = \text{Span}_Q \left\{ e_1 - \frac{2}{b_n} e_{n+1}, e_2, e_3, \ldots, e_{g+1} \right\}.
\]

Suppose first that \(b_n\) is odd. Then we have the following integral basis for \(\mathbb{Z}^{n+1}\):

\[
\left\{ b_n e_1 - 2e_{n+1}, e_2, e_3, \ldots, e_{g+1}, \frac{b_n + 1}{2} e_1 - e_{n+1}, e_{g+2}, \ldots, e_n \right\}.
\]

In terms of this basis, the top right \((g+1) \times (g+2)\) blocks of \(A\) and \(B\) have the following form:

\[
A^{\text{top}} = \begin{pmatrix}
-1 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 1 \\
: & : & : & \cdots & : \\
: & : & 0 & 1 & : \\
0 & 0 & m & \cdots & 0 \\
\end{pmatrix}, \quad B^{\text{top}} = \begin{pmatrix}
0 & & & & b_n \\
0 & & & & 1 \\
& & & & \vdots \\
& & & & \vdots \\
0 & & & & 1 \\
\end{pmatrix}.
\]

It is then easy to check that \(|Q|(A, B, \Lambda_2) = |b_n|m\) and \(|q|(A, B, \Lambda_2, \Lambda_2') = m\).

When \(b_n\) is even, we have the following integral basis

\[
\left\{ \frac{b_n}{2} e_1 - e_n, e_2, e_3, \ldots, e_{g+1}, (b_n + 1)e_1 - 2e_{n+1}, e_{g+2}, \ldots, e_n \right\}.
\]

In terms of this basis, the top right \((g+1) \times (g+2)\) blocks of \(A\) and \(B\) have the same form as in (13). Hence the \(|Q|\)- and \(|q|\)-invariants are as stated in the proposition. □
4 A uniformity estimate for odd degree polynomials

Throughout this section, we fix an odd integer $n = 2g + 1$ with $g \geq 1$. Our goal is to prove Theorem 5(b) by obtaining a bound on the number of integral binary $n$-ic forms having bounded height and discriminant weakly divisible by the square of a large squarefree integer.

Let $m > 0$ be an odd squarefree integer. Recall that we defined a map $\sigma_m : W_m^{(2)} \to W_0(\mathbb{Z})$ in Theorem 3.7 with the following two properties: $f_{\sigma_m(f)} = f$ for every $f \in W_m^{(2)}$, and $|Q|(\sigma_m(f)) = m$. Moreover, in Proposition 3.8, we proved that when $f \in W_m^{(2)}$ is irreducible, it is possible to naturally extend the definition of the $|Q|$-invariant to the set $\text{SL}_n(\mathbb{Z}) \cdot \sigma_m(f)$.

Let $W(\mathbb{Z})^{\text{dist}}$ denote the set of distinguished elements in $W(\mathbb{Z})$, and for any set $L \subset W(\mathbb{Z})$, let $I^{\text{irr}}$ denote the set of elements $w \in L$ such that $f_w$ is irreducible. There is a natural extension of the $|Q|$-invariant to the set $W(\mathbb{Z})^{\text{dist}, \text{irr}}$. For a positive real number $M$ and any set $S \subset W(\mathbb{Z})^{\text{dist}, \text{irr}}$, let $S_{|Q| > M}$ denote the set of elements $w \in S$ with $|Q(w)| > M$. By [5], the number of reducible elements $f \in V(\mathbb{Z})$ with $H(f) < X$ is $O(X^n)$. Hence we have the bound

$$\# \sum_{m > M} \{ f \in W_m^{(2)} : H(f) < X \} \ll \#(\text{SL}_n(\mathbb{Z}) \backslash \{ w \in W(\mathbb{Z})^{\text{dist}, \text{irr}} : H(w) < X \}) + O(X^n). \quad (14)$$

In this section, we obtain an upper bound on the number of $\text{SL}_n(\mathbb{Z})$-orbits on $W(\mathbb{Z})^{\text{dist}, \text{irr}}$ with height bounded by $X$. First, in §4.1, we lay out the reduction theory necessary to express the number of such orbits in terms of the counts of lattice points in certain bounded regions. Then in §4.2, we partition these regions into three parts, the main body, the shallow cusp, and the deep cusp. We prove the desired estimate for each of these parts, thereby obtaining Theorem 5(b).

4.1 Reduction theory and averaging over fundamental domains

Recall that the Iwasawa decomposition of $\text{SL}_n(\mathbb{R})$ is given by $\text{SL}_n(\mathbb{R}) = NTK$, where $N$ is the group of unipotent lower triangular matrices in $\text{SL}_n(\mathbb{R})$, $K = \text{SO}(n)$ is a maximal compact subgroup of $\text{SL}_n(\mathbb{R})$, and $T$ is the split torus of $\text{SL}_n(\mathbb{R})$ consisting of $n \times n$ diagonal matrices with positive diagonal entries and determinant 1. We denote elements in $T$ by $s = \text{diag}(t_1^{-1}, t_2^{-1}, \ldots, t_n^{-1})$, where $t_i > 0$ for $1 \leq i \leq n$ and $t_1 t_2 \cdots t_n = 1$. It will be convenient to make the following change of variables. For $1 \leq i \leq n - 1$, set $s_i$ to be

$$s_i = (t_i/t_{i+1})^{1/n}, \quad \text{which implies} \quad t_i = \prod_{k=1}^{i-1} s_k^{-1} \prod_{k=i}^{n-1} s_k^{n-k}$$

for $1 \leq i < n$. The Haar measure of $G(\mathbb{R})$ in these coordinates is then given by

$$dg = dn\delta(s) d^x dk, \quad \text{where} \quad \delta(s) = \prod_{1 \leq i < j \leq n} \frac{t_j}{t_i} = \prod_{i=1}^{n-1} s_k^{-nk(n-k)}.$$ 

dn and dk are Haar measures on $N$ and $K$, respectively, and $d^x s = \prod_{i=1}^{n-1} s_i^{-1} ds_i$.

We denote the coordinates on $W$ by $a_{ij}, b_{ij}$ for $1 \leq i \leq j \leq n$. These coordinates are eigenvectors for the action of $T$ on the dual $W^*$ of $W$. Denote the $T$-weight of a coordinate $\alpha$ on $W$, or more generally a product $\alpha$ of powers of such coordinates, by $w(\alpha)$. Then $w(a_{ij}) = w(b_{ij}) = t_i^{-1} t_j^{-1}$. It will be useful in what follows to compute the weight of the $Q$-invariant, which is
a homogeneous polynomial of degree $g(g+1)$ in the coordinates of $W_0$. We view the torus $T$ as sitting inside $G_0$. Then by (5), we have

$$w(Q) = \prod_{k=1}^{g} t_k^{-1}. \quad (15)$$

Let $\mathcal{F}$ be a fundamental set for the action of $\text{SL}_n(\mathbb{Z})$ on $\text{SL}_n(\mathbb{R})$ that is contained in a Siegel set, i.e., contained in $N'T'K$, where $N'$ is a set consisting of elements in $N$ whose coefficients are absolutely bounded and $T' \subset T$ consists of elements in $s \in T$ with $s_i \geq c$ for some positive constant $c$. Let $\mathcal{W}(1)$ denote the subset of real binary $n$-ic forms of height bounded by 1 and let $R' = \sigma_1(\mathcal{W}(1))$, where $\sigma_1$ is as in §3.4. Set $R := \mathbb{R}_{>0} \cdot R'$, and note that every distinguished element of $W(\mathbb{R})$ is $\text{SL}_n(\mathbb{R})$-equivalent to some element in $R$.

Let $H_0$ be a nonempty open bounded left $K$-invariant set in $\text{SL}_n(\mathbb{R})$. Denote the set $H_0 \cdot R'$ by $\mathcal{B}_1$. Then $\mathcal{B}_1$ is an absolutely bounded set in $W(\mathbb{R})$. Let $\mathcal{L}$ be any $\text{SL}_n(\mathbb{Z})$-invariant subset of $W(\mathbb{Z})$ consisting of elements that are distinguished over $\mathbb{R}$, and denote the set of elements in $\mathcal{L}$ with height less than $X$ by $\mathcal{L}_X$. Throughout this section, let $Y = X^{1/n}$. Then the averaging method as described in [9, §2.3] yields the bound

$$\#(\text{SL}_n(\mathbb{Z}) \setminus \mathcal{L}_X) \ll \int_{\gamma \in \mathcal{F}} \#(\gamma(Y \mathcal{B}_1 \cap \mathcal{L})) d\gamma \ll \int_{s = (s_i)} \#(s(Y \mathcal{B} \cap \mathcal{L})) d^\mathbb{R} s \quad (16)$$

for some absolutely bounded open set $\mathcal{B}$ containing $\mathcal{B}_1$.

We denote the second integral on the right hand side of (16) by $\mathcal{I}_\mathcal{X}(\mathcal{L})$, and break it up into an integral over the main body, the shallow cusp, and the deep cusp. We define the main body to be the range of the integral where $|a_{11}| \geq 1$ for some element in $s(Y \mathcal{B})$, and denote the main-body portion of $\mathcal{I}_\mathcal{X}(\mathcal{L})$ by $\mathcal{I}_{\mathcal{X}}^{\text{main}}(\mathcal{L})$. We define the shallow cusp to be the range of the integral where $|a_{11}| < 1$ for all elements in $s(Y \mathcal{B})$ but $|a_{ij}| \geq 1$ for some $i, j \leq g$, and denote the shallow-cusp portion of $\mathcal{I}_\mathcal{X}(\mathcal{L})$ by $\mathcal{I}_{\mathcal{X}}^{\text{scusp}}(\mathcal{L})$. We define the deep cusp to be the range of the integral where $|a_{ij}| < 1$ for all $i, j \leq g$ and all elements in $s(Y \mathcal{B})$, and denote the deep-cusp portion of $\mathcal{I}_\mathcal{X}(\mathcal{L})$ by $\mathcal{I}_{\mathcal{X}}^{\text{dcusp}}(\mathcal{L})$. Then

$$\mathcal{I}_\mathcal{X}(\mathcal{L}) = \mathcal{I}_{\mathcal{X}}^{\text{main}}(\mathcal{L}) + \mathcal{I}_{\mathcal{X}}^{\text{scusp}}(\mathcal{L}) + \mathcal{I}_{\mathcal{X}}^{\text{dcusp}}(\mathcal{L}). \quad (17)$$

In the next subsection, we prove bounds for the main body, the shallow cusp, and the deep cusp when $\mathcal{L} = W(\mathbb{Z})_{|\mathcal{Q}|>M}^{\text{dist,irr}}$.

### 4.2 The number of orbits of distinguished elements with large $Q$-invariant

In this subsection, we obtain the following upper bound on $\mathcal{I}_\mathcal{X}(W(\mathbb{Z})_{|\mathcal{Q}|>M}^{\text{dist,irr}})$, thus yielding the same bound on the quantity $\#(\text{SL}_n(\mathbb{Z}) \setminus \{w \in W(\mathbb{Z})_{|\mathcal{Q}|>M}^{\text{dist,irr}} : H(w) < X\})$ by (16).

**Theorem 4.1** We have $\mathcal{I}_\mathcal{X}(W(\mathbb{Z})_{|\mathcal{Q}|>M}^{\text{dist,irr}}) \ll X^{n+1-1/5n+\epsilon} + X^{n+1+\epsilon} / M$.

Note that (14), (16), and Theorem 4.1 immediately imply Part (b) of Theorem 5.

We bound $\mathcal{I}_\mathcal{X}(W(\mathbb{Z})_{|\mathcal{Q}|>M}^{\text{dist,irr}})$ by obtaining bounds for the main body, the shallow cusp, and the deep cusp. We consider first the main body.

**Proposition 4.2** We have $\mathcal{I}_{\mathcal{X}}^{\text{main}}(W(\mathbb{Z})^{\text{dist}}) \ll X^{n+1-1/(5n)+\epsilon}$.

**Proof:** In [17, Proposition 4.6], an upper bound of $o(X^{n+1})$ is obtained on $\mathcal{I}_{\mathcal{X}}^{\text{main}}(W(\mathbb{Z})^{\text{dist}})$. This is proved using the following two ingredients: estimates with a power saving error term on $\mathcal{I}_{\mathcal{X}}^{\text{main}}(\mathcal{L})$ for lattices $\mathcal{L} \subset W(\mathbb{Z})$, and a proof that the density of elements in $W(\mathbb{F}_p)$ that are not $\mathbb{F}_p$-distinguished
is bounded away from 1, independent of $p$. Proposition 4.2 follows from these two ingredients along with an application of the Selberg sieve used exactly as in [28]. □

Next, a bound on the shallow cusp follows directly from the proof of [17, Proposition 4.3]:

**Proposition 4.3** We have $I_{x}^{\text{cusp}}(W(\mathbb{Z})) \ll X^{n+1-1/n}$.

In [17, Proposition 4.3], the shallow and deep cusps were treated simultaneously, but the points in the deep cusp were ruled out since only nondistinguished elements were counted there. Hence the proof of [17, Proposition 4.3] yields the claimed bound in Proposition 4.3.

Finally, to treat the deep cusp, let $U = \{a_{ij}, b_{ij} : 1 \leq i \leq j \leq n\}$ denote the set of coordinates on $W$, and let $U_{0} = \{a_{ij}, b_{ij} | i \leq j, j \geq g + 1\}$ denote the set of coordinates on $W_{0}$. We define a partial order $\preceq$ on $U$ by setting $\alpha \preceq \beta$ if all the powers of $s_{i}$ in $w(\alpha)^{-1}w(\beta)$ are nonnegative. Explicitly, $a_{ij} \preceq a_{i'j'}$ if and only if $i \leq i'$ and $j \leq j'$ (and similarly for $b_{ij}$, as $a_{ij}$ and $b_{ij}$ have the same weight). A subset $Z$ of $U_{0}$ is saturated if for any $\beta \in Z$ and any $\alpha \in U_{0}$ with $\alpha \preceq \beta$, the coordinate $\alpha$ also lies in $Z$. We pick positive constants $c_{ij}$ for $1 \leq i \leq j \leq n$ such that:

(a) If $|Yw(a_{ij})| < c_{ij}$, then $|a_{ij}| < 1$ and $|b_{ij}| < 1$ for every $(A, B) \in s(YB)$.

(b) For all $s \in T'$ and $a_{ij} \preceq a_{i'j'}$, we have $w(a_{ij})/c_{ij} \leq w(a_{i'j'})/c_{i'j'}$.

More explicitly, we may choose $c_{nn}$ to be sufficiently small and take

$$c_{ij} = \left(\sup_{s \in T'} \frac{w(a_{ij})}{w(a_{nn})}\right) c_{nn}, \quad \text{for } i \leq j \leq n.$$  

The significance of these constants $c_{ij}$ is the following: for every $Y > 1$, first, if $Yw(a_{ij}) < c_{ij}$, then every integral element in $s(YB)$ has $a_{ij}$- and $b_{ij}$-coordinates equal to 0; and second, if $a_{ij} \preceq a_{i'j'}$, then $Yw(a_{i'j'}) < c_{i'j'}$ implies $Yw(a_{ij}) < c_{ij}$.

We have the following result due to Davenport which we use to estimate the number of lattice points in bounded regions.

**Proposition 4.4 ([12])** Let $\mathcal{R}$ be a bounded, semi-algebraic multiset in $\mathbb{R}^{n}$ having maximum multiplicity $m$ that is defined by at most $k$ polynomial inequalities, each having degree at most $\ell$. Let $\mathcal{R}'$ denote the image of $\mathcal{R}$ under any (upper or lower) triangular, unipotent transformation of $\mathbb{R}^{n}$. Then the number of lattice points (counted with multiplicity) contained in the region $\mathcal{R}'$ is given by

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

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

The following lemma gives conditions that ensure an element in $W(\mathbb{R})$ has discriminant 0.

**Lemma 4.5** Suppose that $(A, B) \in W(\mathbb{R})$ satisfies $a_{ij} = b_{ij} = 0$ for all $i \leq k$ and $j \leq n-k$ for some $k \in \{1, \ldots, g\}$. Then the discriminant of $(A, B)$ is 0.

**Proof:** One checks that $f_{A, B}$ has a square factor of degree $k$ and so has discriminant 0. □

The next lemma states that when $\mathcal{L} \subset W(\mathbb{Z})$ consists of elements with nonzero discriminant, the integral defining $I_{X}(\mathcal{L})$ can be cut off by conditions of the form $s_{i} \ll X^{\Theta}$ for some absolute constant $\Theta$ depending only on $n$.  

18
Lemma 4.6 There exists an absolute constant Θ depending only on n such that if \( s \in T' \) with \( s_i \gg X^\Theta \) for some \( i \), then \( s(YB) \cap W(\mathbb{Z}) \) contains only points with discriminant 0.

Proof: Let \( s = \text{diag}(t_1, \ldots, t_n) \in T' \); then \( t_1 \gg t_2 \gg \cdots \gg t_n \) and \( t_1 t_2 \cdots t_n = 1 \). Because of the relation between the \( t_j \)'s and the \( s_i \)'s, it suffices to prove that if \( s(YB) \) contains an integral element with nonzero discriminant, then \( t_1 \) is bounded from above by some power of \( X \) or, equivalently, \( t_n \) is bounded from below by some power of \( X \). By Lemma 4.5, for \( s(YB) \) to contain an integral element with nonzero discriminant, we must have \( Yw(a_{k,n-k}) \gg 1 \) for every \( k \in \{1, \ldots, g\} \). That is, \( t_k t_{n-k} \ll Y \) for every \( k \in \{1, \ldots, g\} \). Multiplying these conditions together, we obtain \( t_n \gg Y^{-g} \).

We now estimate the contribution to \( I_X(W(\mathbb{Z})_{|Q|>M}^{\text{dist.,irr}}) \) coming from the deep cusp.

Proposition 4.7 We have \( \mathcal{I}_X(W(\mathbb{Z})_{|Q|>M}^{\text{cusp}}) \ll_{\epsilon} X^{n+1+\epsilon/M} \).

Proof: For a subset \( Z \) of \( U_0 \), let \( T'_Z \) denote the subset of \( s \in T' \) with \( Y^{g(g+1)}w(Q) \gg M \), and \( |Yw(a_{ij})| < c_{ij} \) precisely for those \((i, j)\) where \( a_{ij} \in Z \) or \( b_{ij} \in Z \). Note that \( T'_Z \) is empty if \( Z \) is not saturated. Define

\[
N(Z, X) := \int_{s \in T'_Z} \#(s(YB) \cap W_0(\mathbb{Z})) \delta(s) d^X s
\]

\[
\ll \int_{s \in T'_Z} \left( \prod_{a \in U_0 \backslash Z} Yw(\alpha) \right) \delta(s) d^X s
\]

\[
= \int_{s \in T'_Z} \left( \prod_{a \in U_0} Yw(\alpha) \right) \left( \prod_{a \in Z} Y^{-1}w(\alpha)^{-1} \right) \delta(s) d^X s,
\]

where the bound on the second line follows from Proposition 4.4. Let \( U' := \{a_{ij}, b_{ij} \mid i + j < n\} \). If \( Z \) is saturated and not contained in \( U' \), then \( Z \) contains \( a_{k,n-k} \) for some \( k = 1, \ldots, g \). Hence, for any \( s \in T'_Z \), every integral element in \( s(YB) \cap W_0(\mathbb{Z}) \) satisfies \( a_{ij} = b_{ij} = 0 \), for \( i \leq k \) and \( j \leq n - k \), and so has zero discriminant by Lemma 4.5. Therefore,

\[
\mathcal{I}_X(W(\mathbb{Z})_{|Q|>M}^{\text{cusp}}) \ll \sum_Z N(Z, X)
\]

where the sum is over saturated subsets \( Z \) of \( U_0 \) contained in \( U' \).

Now

\[
\prod_{a \in U_0} Yw(\alpha) = Y^{n(n+1)-g(g+1)}(t_1 \cdots t_{g+1})^{2g+4}
\]

\[
= \frac{Y^{n(n+1)}}{Y^{g(g+1)}w(Q)}(t_1 \cdots t_{g+1})^{g+2}(t_{g+2} \cdots t_n)^{-g+1}
\]

\[
= \frac{Y^{n(n+1)}}{Y^{g(g+1)}w(Q)} \prod_{i=1}^{g+1} \prod_{j=g+2}^{n} \frac{t_i}{t_j}.
\]

(18)

Fix a saturated subset \( Z \) of \( U_0 \) contained in \( U' \). We define a map \( \pi : Z \to U_0 \backslash U' \) by

\[
\pi(a_{ij}) = a_{i,n-i}, \quad \pi(b_{ij}) = b_{n-j,i}.
\]

Note that for any \( \alpha \in Z \), we have \( \pi(\alpha) \notin U' \) and so \( Yw(\pi(\alpha)) \gg 1 \). Furthermore, for every \( \alpha \in U' \), we have \( \alpha \preceq \pi(\alpha) \) and so \( w(\pi(\alpha))/w(\alpha) \gg 1 \). Hence for any \( s \in T'_Z \),

\[
\prod_{a \in Z} (Yw(\alpha))^{-1} \ll \prod_{a \in Z} Yw(\pi(\alpha))/Yw(\alpha) = \prod_{g+2 \leq i < j \leq n} \frac{t_i}{t_j} \prod_{1 \leq i < j \leq g+1} \frac{t_i}{t_j}.
\]

(19)
Here the first product on the right hand side is the contribution from all \( a_{ij} \in U_1 \), and the second product is the contribution from all \( b_{ij} \in U_1 \). Combining (18) and (19) gives, for any \( s \in T' \),

\[
\prod_{\alpha \in \mathcal{T}(\delta_0), \mathcal{Z}} Y^w(\alpha) \ll \frac{Y^{n(n+1)}}{Y^g(1+1)w(Q)} \prod_{1 \leq i < j \leq n} \frac{t_i}{t_j} = \frac{Y^{n(n+1)}}{Y^g(1+1)w(Q)} \delta(s)^{-1} \ll \frac{Y^{n(n+1)}}{M} \delta(s)^{-1}.
\]

Since each \( s_i \) is bounded below by an absolute constant and bounded above by a power of \( X \) by Lemma 4.6, we obtain

\[
N(\mathcal{Z}, X) = O\left(\frac{X^{n+1+\epsilon}}{M}\right).
\]

The proof is completed by summing over all saturated subsets \( \mathcal{Z} \) contained in \( U_1 \). \( \square \)

5 A bound on the number of singular symmetric matrices in skewed boxes

Let \( n \geq 2 \) be a positive integer and let \( S = \text{Sym}_2(n) \) denote the space of symmetric \( n \times n \) matrices. Let \( | \cdot | \) denote Euclidean length on \( S(\mathbb{R}) \) obtained by identifying \( S(\mathbb{R}) \) with \( \mathbb{R}^{\dim S} = \mathbb{R}^{n(n+1)/2} \). Let \( D \subset S(\mathbb{R}) \) be a bounded open set. For an integer \( r \) with \( 1 \leq r < n \), let \( S(\mathbb{Z})_r \) denote the set of elements in \( S(\mathbb{Z}) \) having rank \( r \). In [15], Eskin and Katznelson obtained asymptotics for the number of elements in \( YD \cap S(\mathbb{Z})_r \) for \( r \in \{1, \ldots, n-1\} \).

In this paper, we will not need exact asymptotics; upper bounds will suffice. In this section, our goal is to obtain upper bounds on the number of elements of \( S(\mathbb{Z})_r \) in skew balls.

The group \( \text{SL}_n \) acts on \( S \) via \( \gamma(A) = \gamma A \gamma^t \) for \( \gamma \in \text{SL}_n \) and \( A \in S \). Let \( T \subset \text{SL}_n(\mathbb{R}) \) denote the subgroup of diagonal matrices with positive coefficients. We denote elements in \( T \) by \( s = \text{diag}(t_1^{-1}, \ldots, t_n^{-1}) \). We are interested in studying the skew ball \( s(YD) \). By symmetry, we may assume that \( s \in T' \), i.e., we have \( t_1 \gg t_2 \gg \ldots \gg t_n \). Moreover, in light of Lemma 4.6, we will assume that \( t_1 \ll Y^\Theta \) and \( t_n \gg Y^{-\Theta} \) for some absolute constant \( \Theta \) depending only on \( n \).

For \( s \in T \) and \( r \in \{1, \ldots, n-1\} \), we define the constants \( C(r, s) \) by

\[
C(r, s) = \prod_{i=1}^{n-r} \prod_{j=1}^{n-i} \frac{t_j}{t_{n-i+1}} = \prod_{1 \leq i < j \leq n-r} \frac{t_i}{t_j}.
\]

When \( s \in T' \), these constants satisfy

\[
C(r, s) \ll C(n-1, s) = \prod_{1 \leq i < j \leq n} \frac{t_i}{t_j} = \delta(s)^{-1},
\]

where as before \( \delta(s) \) is the character of the torus appearing in the Haar measure of \( \text{SL}_n(\mathbb{R}) \).

Finally, for \( 1 \leq r < n \), a positive real number \( Y \), and \( s \in T \), let \( N_r(Y, s) \) denote the number of elements in \( s(YD) \cap S(\mathbb{Z})_r \). We prove the following result.

**Theorem 5.1** Let \( n \geq 2 \) and \( 1 \leq r < n \) be positive integers. Let \( \Theta > 0 \) be a real number. Let \( Y > 1 \) be a real number, and let \( s \in T' \) with \( t_1 \ll Y^\Theta \) and \( t_n \gg Y^{-\Theta} \). Then

\[
N_r(Y, s) = O(C(r, s)Y^{nr/2} \log Y)
\]

where the implied constants depend only on \( n \) and \( \Theta \).
The case $s = 1$ of Theorem 5.1 follows from the work of Eskin-Katznelson [15]. Their strategy is to express the set of singular symmetric matrices of rank $r$ as a union of lattices, each of which consists of elements having a fixed row span. They count the number of elements in each such lattice having bounded norm, and then sum over all possible row spans. We follow this strategy, explaining the modifications necessary to bound integer points in skew balls.

Fix positive integers $k$ and $m$ with $k \leq m$, and a lattice $\Lambda$ in $\mathbb{R}^n$ of rank $r$. A basis $\{\ell_1, \ldots, \ell_k\}$ of $\Lambda$ is reduced if the product $|\ell_1||\ell_2| \cdots |\ell_k|$ is minimal among all integral bases of $\Lambda$. It is almost reduced if

$$|\ell_1||\ell_2| \cdots |\ell_k| \ll d(\Lambda),$$

where $d(\Lambda)$ denotes the covolume of $\Lambda$ in $\Lambda \otimes \mathbb{R}$, and the implied constant in the inequality depends only on $n$. If we order an almost reduced basis $\{\ell_1, \ldots, \ell_k\}$ by length, then the $i$-th successive minimum of $\Lambda$ is within a constant multiple (depending only on $n$) of $|\ell_i|$ for every $i = 1, \ldots, k$. To bound the number elements of $\Lambda$ in a ball, we use the following result of Schmidt [26].

**Proposition 5.2** Let $\Lambda$ be a rank $k$ lattice in $\mathbb{R}^m$ and $D$ a bounded open domain in $\mathbb{R}^m$. Let $\mu_1, \ldots, \mu_k$ be the successive minima of $\Lambda$. Then for $Y > 0$, we have

$$\#(YD \cap \Lambda) = O\left(\max_{1 \leq j \leq k} \frac{Y^j}{\mu_1 \cdots \mu_j}\right). \quad (21)$$

We use the notation of Theorem 5.1. Given a lattice $\Lambda \subset \mathbb{Z}^n$ of rank $r$, let $S(\Lambda)$ denote the lattice of symmetric matrices $B \in S(\mathbb{Z})$ such that the row space (equivalently, the column space) of $B$ is contained in $\Lambda \otimes \mathbb{R}$. For two vectors $v_1$ and $v_2$ in $\mathbb{R}^n$, we define

$$v_1 * v_2 := \begin{cases} v_1 \cdot v_2^1 + v_2 \cdot v_1^1 & \text{if } v_1 \text{ and } v_2 \text{ are linearly independent;} \\ v_1 \cdot v_2^2 & \text{otherwise.} \end{cases} \quad (22)$$

Then $v_1 * v_2 = v_2 * v_1 \in S(\text{Span}\{v_1, v_2\})$, and

$$|v||w| \leq |v * w| \leq 2|v||w|. \quad (23)$$

Fix $\gamma \in \text{SL}_n(\mathbb{R})$. (For our applications, we will take $\gamma \in T$.) Let $\Lambda \subset \mathbb{Z}^n$ be a primitive lattice of rank $r$. We bound the number of elements in $\gamma^{-1}(YD) \cap S(\Lambda)$ using the bijection

$$\gamma^{-1}(YD) \cap S(\Lambda) \rightarrow YD \cap \gamma(S(\Lambda))$$

$$A \mapsto \gamma A \gamma^{-1}$$

and instead bounding the number of elements in $YD \cap \gamma(S(\Lambda))$. We thus study the lattice $\gamma(S(\Lambda)) \subset S(\mathbb{R})$. The next result, which gives an almost reduced basis for $\gamma(S(\Lambda))$ in terms of an almost reduced basis of $\gamma \Lambda$, follows from the proofs of [15, Proposition 3.3] and [15, Lemma 3.5].

**Theorem 5.3** Fix $\gamma \in \text{SL}_n(\mathbb{R})$. Let $\Lambda \subset \mathbb{Z}^n$ be a primitive lattice of rank $r$, and let $\{\ell_1, \ldots, \ell_r\}$ be a basis for $\gamma \Lambda$. Then $\{\ell_i * \ell_j : 1 \leq i \leq j \leq r\}$ is a basis for $\gamma(S(\Lambda))$. Furthermore,

$$d(\gamma(S(\Lambda))) = 2^{(r-1)/2}d(\gamma \Lambda)^{r+1} = 2^{(r-1)/4}d(\Lambda)^{r+1}.$$ 

In particular, if $\{\ell_1, \ldots, \ell_r\}$ is almost reduced, then so is $\{\ell_i * \ell_j : 1 \leq i \leq j \leq r\}$.

Next, by the proof of [15, Lemma 4.1], we have the following result giving a necessary condition for the set $YD \cap \gamma(S(\Lambda))$ to be nonempty.
Proposition 5.4 Let \( \gamma \in \text{SL}_n(\mathbb{R}) \) and let \( \Lambda \subset \mathbb{Z}^n \) be a primitive lattice of rank \( r \) such that the successive minima of \( \gamma \Lambda \) are \( \mu_1 \leq \ldots \leq \mu_r \). If \( \#(YD \cap \gamma(S(\Lambda))) > 0 \), then \( \mu_i \mu_j \leq c_1 Y \) for every pair \( (i, j) \) with \( i + j \leq r + 1 \), for some constant \( c_1 \) depending only on \( n \).

We now prove an upper bound on \( \#(YD \cap \gamma(S(\Lambda))) \).

Proposition 5.5 Let \( \gamma \in \text{SL}_n(\mathbb{R}) \) and let \( \Lambda \subset \mathbb{Z}^n \) be a primitive lattice of rank \( r \) such that the successive minima of \( \gamma \Lambda \) are \( \mu_1 \leq \ldots \leq \mu_r \). Then

\[
\#(YD \cap \gamma(S(\Lambda))) = O\left(\frac{Y^{r(r+1)/2}}{(d(\Lambda))^{r+1}} \prod_{1 \leq i < j \leq r} \frac{\mu_j}{\mu_i} \right).
\]

(24)

Proof: Let \( U(r) \) denote the set of pairs \( (i, j) \) of positive integers such that \( i \leq j \leq r \) and \( i + j > r + 1 \). In other words, elements in \( U(r) \) correspond to the successive minima of the lattice \( \gamma(S(\Lambda)) \) that are \( \gg Y \). By Proposition 5.2, Theorem 5.3, and (23), we have

\[
\#(YD \cap \gamma(S(\Lambda))) \leq \frac{Y^{r(r+1)/2}}{d(\Lambda)^{r+1}} \prod_{(i, j) \in U(r)} \left( \frac{Y}{\mu_i \mu_j} \right)^{i-1}.
\]

Assume that \( \#(YD \cap \gamma(S(\Lambda))) > 0 \). Then \( \mu_{r+1-j} \mu_j \leq Y \) for all \( 1 \leq j \leq r \) by Proposition 5.4. Thus

\[
\#(YD \cap \gamma(S(\Lambda))) \leq \frac{Y^{r(r+1)/2}}{d(\Lambda)^{r+1}} \prod_{(i, j) \in U(r)} \left( \frac{Y}{\mu_i \mu_j} \right)^{i-1} \frac{Y}{\mu_{r+1-j} \mu_j}.
\]

(25)

Since \( i \leq j \) and \( i + j > r + 1 \) for \( (i, j) \) \in \( U(r) \), we have the following injection:

\[
U(r) \rightarrow \{(k, \ell) : 1 \leq k < \ell \leq r : k + \ell \leq r + 1\}
\]

(26)

\[
(i, j) \mapsto (r + 1 - j, i).
\]

Since \( \frac{\mu_r}{\mu_i} \geq 1 \) for \( j > i \), the injection (26) implies that the product of the ratios \( \mu_j/\mu_i \) in (24) is at least as large as the product of the ratios \( \mu_i/\mu_{r+1-j} \) in (25). The result follows. \( \square \)

We now sum over the appropriate lattices \( \Lambda \subset \mathbb{Z}^n \) having rank \( r \). To this end, we fix an element \( s = \text{diag}(t_1^{-1}, t_2^{-1}, \ldots, t_n^{-1}) \in T' \). We will apply the previous results with \( \gamma = s^{-1} \). Set \( L = (L_1, \ldots, L_r) \) with \( 0 < L_1 \leq L_2 \leq \cdots \leq L_r \). Let \( \Sigma(L, s) \) denote the set of primitive lattices \( \Lambda \subset \mathbb{Z}^n \) of rank \( r \) whose successive minima \( \mu_1, \ldots, \mu_r \) of \( s^{-1} \Lambda \) satisfy \( L_i \leq \mu_i < 2L_i \) for each \( i \).

Lemma 5.6 Let \( L = (L_1, \ldots, L_r) \) and \( s = \text{diag}(t_1^{-1}, \ldots, t_n^{-1}) \in T' \). Then there is a constant \( c' > 0 \) depending only on \( n \) such that if \( \# \Sigma(L, s) > 0 \), then \( L_it_j^{-1} > c' \) for all \( (i, j) \) with \( i + j \geq n + 1 \).

Proof: Since \( \# \Sigma(L, s) > 0 \), there exists an integral lattice \( \Lambda \subset \mathbb{Z}^n \) of rank \( r \) with basis \( \{\ell_1, \ldots, \ell_r\} \) such that \( |s^{-1} \ell_i| < 2L_i \) for \( i \in \{1, \ldots, r\} \). For \( 1 \leq j \leq n \), let \( u_{ij} \) denote the (integral) \( j \)-th entry of \( \ell_i \). Then \( |u_{ij}| \leq 2L_i t_j^{-1} \) for every \( 1 \leq i \leq r \) and \( 1 \leq j \leq n \). The assumption that \( s \in T' \) implies that \( L_it_j^{-1} < L_i t_j^{-1} \) whenever \( i \leq i' \) and \( j \leq j' \).

Suppose that there is an integer \( k \) with \( 1 \leq k \leq r \) such that \( L_k t_{n+1-k}^{-1} > c'' \) for some sufficiently small constant \( c'' > 0 \). Then \( |u_{ij}| < 1 \), and thus \( u_{ij} = 0 \) for all \( (i, j) \) with \( 1 \leq i \leq k \) and
1 ≤ j ≤ n + 1 − k. However, this implies that the vectors ℓ₁, . . . , ℓₖ are not linearly independent, a contradiction. Hence such a k does not exist and Lₖtᵦ−¹ₙ₊₁₋ₖ ≥ 1 for all k, implying the result. □

We now determine an upper bound for Σ(L, s).

**Proposition 5.7** Let L = (L₁, . . . , Lᵦ) and s = diag(t⁻¹₁, . . . , t⁻¹ᵦ) ∈ T'. Then

\[
#Σ(L, s) = O\left(\left(\prod_{1 ≤ i ≤ j ≤ r} \frac{L_i}{L_j}\right)C(r, s)\right)
\]

where C(r, s) is defined as in (20).

**Proof:** We count lattices Λ by counting r-tuples of vectors (ℓ₁, . . . , ℓᵦ) such that each ℓᵢ ∈ s⁻¹Zⁿ satisfies Lᵢ ≤ |ℓᵢ| < 2Lᵢ and such that {ℓ₁, . . . , ℓᵦ} is a reduced basis of the lattice it generates. For each i = 1, . . . , r, let α(i) be the largest integer such that Lᵢtᵦ−¹ᵢ ≤ α', where α' is as in Lemma 5.6, or let α(i) = 0 if no such integer exists. By Proposition 5.2, the number of possibilities for ℓᵢ is

\[
\ll \prod_{j=1}^{n} \max(L_i t_j⁻¹, 1) \ll L_i^n \prod_{j=1}^{α(i)} (L_i t_j⁻¹).
\]

However, once ℓ₁ is fixed, and given a vector ℓ₂, at most two of ℓ₂ − kℓ₁ can be part of a reduced basis for k ∈ Z. Since ≥ L₂/L₁ vectors ℓ₂ − kℓ₁ satisfy the same size bound as ℓ₂ (namely, those with k ≪ L₂/L₁), the number of choices for the pair (ℓ₁, ℓ₂) that are part of a reduced basis is

\[
\frac{L₁}{L₂}L₁^n L₂^n \prod_{j=1}^{α(1)} (L₁⁻¹ t_j) \prod_{j=1}^{α(2)} (L₂⁻¹ t_j).
\]

Continuing in this way, we obtain the bound

\[
#Σ(L, s) \ll (L₁ L₂ . . . Lᵦ)\left(\prod_{1 ≤ i < j ≤ r} \frac{L_i}{L_j}\right)\left(\prod_{i=1}^{r} \prod_{j=1}^{α(i)} L_i⁻¹ t_j\right),
\]

(28)

By Lemma 5.6, we have α(i) ≤ n − i for i ∈ {1, . . . , r}. Therefore,

\[
\prod_{i=1}^{r} \prod_{j=1}^{α(i)} L_i⁻¹ t_j \ll \prod_{i=1}^{r} \prod_{j=1}^{α(i)} L_i⁻¹ t_j L_i t_α(i)+1 \ll \prod_{i=1}^{r} \prod_{j=1}^{n-i} t_j \ll \prod_{i=1}^{r} \prod_{j=1}^{n-i} t_{n-i+1} = C(r, s).
\]

(29)

Equations (28) and (29) yield the desired result. □

We are now ready to prove the main result of this section.

**Proof of Theorem 5.1:** Let L = (L₁, . . . , Lᵦ) be a tuple such that 0 < L₁ ≤ L₂ ≤ . . . ≤ Lᵦ. Then, by Lemma 5.6, Proposition 5.4 and the definition of T', we see that for there to exist a lattice Λ ∈ Σ(L, s) such that #(YD ∩ s⁻¹(S(Λ))) > 0, we must have

\[
Y^{−Θ₁} ≪ L₁ ≤ . . . ≤ Lᵦ ≪ Y^{Θ₂} \quad \text{and} \quad L₁ . . . Lᵦ ≪ Y^{r/2}
\]

for some absolute constants Θ₁, Θ₂ > 0. For any such Λ, Proposition 5.5 states that

\[
#(YD ∩ s⁻¹(S(Λ))) \ll \frac{Y^{r(r+1)/2}}{(L₁ . . . Lᵦ)^{r+1}} \prod_{1 ≤ i < j ≤ r} \frac{L_i}{L_i},
\]

23
Thus
\[ N_r(Y, s) \ll \sum_L \# \Sigma(L, s) \frac{Y^r(r+1)/2}{(L_1 \ldots L_r)^{r+1}} \prod_{1 \leq i < j \leq r \atop i+j \leq r+1} \frac{L_j}{L_i}, \]
where the sum is over \( r \)-tuples \( L = (L_1, \ldots, L_r) \) with \( L_1 \leq L_2 \leq \cdots \leq L_n \) that partition the region \( \{(\mu_1, \ldots, \mu_r) \in [Y^{-\Theta}, Y^{\Theta}]^r : \mu_1 \leq \cdots \leq \mu_r \} \) into dyadic ranges. The sum over \( L \) has length \( O(\log^r Y) \). Using the upper bound on \( \# \Sigma(L, s) \) in Proposition 5.7, we obtain
\[
N_r(Y, s) \ll \sum_L C(r, s)(L_1 \ldots L_r)^n \frac{Y^{r(r+1)/2}}{(L_1 \ldots L_r)^{r+1}} \left( \prod_{1 \leq i < j \leq r \atop i+j \leq r+1} \frac{L_j}{L_i} \right) \left( \prod_{1 \leq i < j \leq r} \frac{L_i}{L_j} \right) \]
\[
\ll C(r, s)Y^{nr/2} \log^r Y.
\]
This concludes the proof of Theorem 5.1. \( \square \)

6 A uniformity estimate for even degree polynomials

We fix an even integer \( n = 2g + 2 \) with \( g \geq 1 \). Our goal is to prove Theorem 5(c) by obtaining a bound on the number of integral binary \( n \)-ic forms having bounded height having discriminant weakly divisible by the square of a large squarefree integer.

Throughout this section, we write \( V := V_n \) and \( W := W_{n+1} \). Let \( m > 0 \) be an odd squarefree integer, and let \( \mathcal{W}^{(2)}_m := \mathcal{W}^{(2)}_{m,n} \). We also define the following auxiliary sets:
\[
V(\mathbb{Z})^{\text{red}} := \{ f \in V(\mathbb{Z}) : \text{Gal}(f(x, 1)) \neq S_n \},
\]
\[
V(\mathbb{Z})^{\Delta \text{small}} := \{ f \in V(\mathbb{Z}) : \Delta(f) < H(f)^{2n-2-\kappa} \},
\]
\[
\mathcal{W}^{(1)}_m := \{ f \in \mathcal{W}^{(2)}_m : m \mid f(0, 1) \},
\]
\[
\mathcal{W}_m^{\text{gen}} := \{ f \in \mathcal{W}^{(2)}_m : \gcd(m, f(0, 1)) = 1 \text{ and } f \not\in V(\mathbb{Z})^{\text{red}} \cup V(\mathbb{Z})^{\Delta \text{small}} \},
\]
where \( \kappa > 0 \) is a small constant (whose exact value will be optimized later) and \( \text{Gal} \) denotes the Galois group. Then, for any \( M > 0 \), we have the following containment:
\[
\bigcup_{m > M \text{ squarefree}} \mathcal{W}^{(2)}_m \subset V(\mathbb{Z})^{\text{red}} \cup V(\mathbb{Z})^{\Delta \text{small}} \cup \bigcup_{m > \sqrt{M} \text{ squarefree}} \mathcal{W}^{(1)}_m \cup \bigcup_{m > \sqrt{M} \text{ squarefree}} \mathcal{W}_m^{\text{gen}}.
\]

The number of elements in \( V(\mathbb{Z})^{\text{red}} \) having height less than \( X \) was bounded by \( O(X^n) \) in [5]. We next prove a bound on the number of elements in \( V(\mathbb{Z})^{\Delta \text{small}} \) of bounded height.

Lemma 6.1 The number of integral binary \( n \)-ic forms with height less than \( X \) and absolute discriminant less than \( X^{2n-2-\kappa} \) is \( O(X^{n+1-\eta}) \).

Proof: Set \( \eta := \kappa/(2n-2) \). The number of integral binary \( n \)-ic forms \( a_0x^n + \cdots + a_ny^n \) with height less than \( X \) such that \( |a_0| \leq X^{1-\eta} \) is \( O(X^{n+1-\eta}) \). Hence we assume \( |a_0| > X^{1-\eta} \).

Now fix integers \( a_0, \ldots, a_{n-1} \) with \( |a_i| \leq X \) and \( |a_0| > X^{1-\eta} \). The discriminant of \( a_0x^n + \cdots + a_ny^n \) is a polynomial \( F(a_n) \) in \( a_n \) of degree \( n-1 \) with leading coefficient \( C_na_0^{n-1} \) for some nonzero constant \( C_n \). Let \( r_1, \ldots, r_{n-1} \in \mathbb{C} \) be the \( n-1 \) roots of \( F(x) \). Then
\[
F(a_n) = C_na_0^{n-1}(a_n - r_1) \cdots (a_n - r_{n-1}).
\]
Since $|F(a_n)| < X^{2n-2-\eta}$, we have $(a_n - r_1) \cdots (a_n - r_{n-1}) \ll X^{n-1-(n-1)\eta}$. Hence $|a_n - r_i| \ll X^{1-\eta}$ for some $i = 1, \ldots, n-1$. The number of such integers $a_n$ is $O(X^{1-\eta})$. Since there are $O(X^n)$ choices for $a_0, \ldots, a_{n-1}$, we obtain the desired bound. □

A direct application of a quantitative version of the Ekedahl sieve as in [3, Theorem 3.3] implies the following bound on the number of elements of bounded height belonging to $\mathcal{W}_m^{(1\#)}$ for large $m$.

**Lemma 6.2** We have
\[
\# \mathbb{P}_{m > \sqrt{M}, m \text{ squarefree}} \{ f \in \mathcal{W}_m^{(1\#)} : H(f) < X \} = O\left( \frac{X^{n+1}}{\sqrt{M}} + X^n \right).
\]

To prove Theorem 5(c), it thus remains to obtain an upper bound for
\[
\# \mathbb{P}_{m > \sqrt{M}, m \text{ squarefree}} \{ f \in \mathcal{W}_m^{\text{gen}} : H(f) < X \} = O\left( \frac{X^{n+1}}{\sqrt{M}} + X^n \right) \tag{35}
\]

In §3.5, we defined a map $\sigma_m$ from the set of elements $f \in \mathcal{W}_m^{(2)}$ with $\gcd(m, f(0,1)) = 1$ to $W(\mathbb{Z})$ such that $f_{\sigma_m(f)} = xf$ and $|q|(\sigma_m(f)) = m$. For any $M > 0$, define the set $\mathcal{L}(M)$ by
\[
\mathcal{L}(M) := \bigcup_{m > M} \text{SL}_{n+1}(\mathbb{Z}) \cdot \sigma_m(\mathcal{W}_m^{\text{gen}}).
\]

Then (35) is \ll
\[
\# \left( \text{SL}_{n+1}(\mathbb{Z}) \setminus \{ f \in \mathcal{L}(M) : H(f) < X \} \right) \ll \mathcal{I}_X(\mathcal{L}(M)) \tag{36}
\]

where
\[
\mathcal{I}_X(\mathcal{L}(M)) = \int_{s \in T'} \#(sYB) \cap \mathcal{L}(M) \delta(s) d^\times s
\]

is as defined immediately after (16), where $Y$ is now taken to be $X^{1/(n+1)}$ throughout this section. Moreover, exactly as in the paragraph leading up to (17), we break up $\mathcal{I}_X(\mathcal{L}(M))$ into three parts—corresponding to the main body, the shallow cusp, and the deep cusp—and again write
\[
\mathcal{I}_X(\mathcal{L}(M)) = \mathcal{T}_X^{\text{main}}(\mathcal{L}(M)) + \mathcal{T}_X^{\text{scusp}}(\mathcal{L}(M)) + \mathcal{T}_X^{\text{dcusp}}(\mathcal{L}(M)).
\]

The rest of this section is dedicated to obtaining an upper bound on $\mathcal{I}_X(\mathcal{L}(M))$. Every element $(A,B) \in \mathcal{L}(M)$ satisfies $\det(B) = 0$ since $f_{A,B}$ is divisible by $x$. In §4, we used vanishing conditions on the coefficients $\{a_{ij}, b_{ij}\}$ of $W$ to estimate the number of integral pairs $(A,B)$ in skewed domains of $W(\mathbb{R})$. Now, since we also need to impose the condition that $B$ has determinant $0$, we use the setup of §5 to count the number of such $B$’s in skewed bounded domains by fibering over the row space of $B$.

In §6.1, we thus further break up the three parts of $\mathcal{I}_X(\mathcal{L}(M))$ into sums over row spaces of the singular matrix $B$. We also obtain some preliminary bounds on $\mathcal{I}_X(\mathcal{L}(M))$, and give some conditions that ensure that a pair $(A,B)$ has discriminant $0$. In §6.2, §6.3, and §6.4, we then prove the desired upper bounds on $\mathcal{T}_X^{\text{main}}(\mathcal{L}(M))$, $\mathcal{T}_X^{\text{scusp}}(\mathcal{L}(M))$, and $\mathcal{T}_X^{\text{dcusp}}(\mathcal{L}(M))$, respectively. In conjunction with (34), (36), and Lemmas 6.1–6.2 this will yield Theorem 5(c).
6.1 Setup and preliminary bounds

Coordinate systems, weight functions, and summing over row spaces

Let $S(\mathbb{Z})$ denote the set of $(n+1) \times (n+1)$ integral symmetric matrices. For any primitive lattice $\Lambda$ of $\mathbb{Z}^{n+1}$, let $S(\Lambda)$ denote the sublattice of $S(\mathbb{Z})$ consisting of elements $B \in S(\mathbb{Z})$ with row space contained in $\Lambda$. For $L = (L_1, \ldots, L_n)$ with $L_i \in \mathbb{R}$ and $L_1 \leq L_2 \leq \cdots \leq L_n$ and $s \in T'$, let $\Sigma(L, s)$ denote the set of primitive lattices $\Lambda \subset \mathbb{Z}^{n+1}$ of rank $n$ such that the successive minima $\mu_1, \ldots, \mu_n$ of $s^{-1}\Lambda$ satisfy $L_1 \leq \mu_i \leq 2L_i$ for each $i$. We define $S(L, s) \subset S(\mathbb{Z})$ by

$$S(L, s) := \bigcup_{\Lambda \in \Sigma(L, s)} S(\Lambda).$$

We next introduce coordinate systems and weight functions. Let $\mathcal{M} := \{\ell_{ij} : 1 \leq i \leq n, 1 \leq j \leq n+1\}$ denote the set of coordinates of $n$-tuples of vectors in $\mathbb{R}^{n+1}$. We define

$$w_L(\ell_{ij}) := L_it_j^{-1}.$$  

The significance of $w_L$ is the following. Let $\Lambda \in \Sigma(L, s)$ be a lattice with an integral basis $\{\ell_1, \ldots, \ell_n\}$ such that $\{s^{-1}\ell_1, \ldots, s^{-1}\ell_n\}$ is a Minkowski-reduced basis for $s^{-1}\Lambda$. Then the $j$th coefficient of $\ell_i$ is $\ll L_it_j^{-1} = w_L(\ell_{ij})$. In particular, for the absolute value of the $j$th coefficient of $\ell_i$ to be nonzero, we must have $w_L(\ell_{ij}) \gg 1$. When $L$ is implicit, we will write $w$ in place of $w_L$.

Let $\mathcal{K}$ denote the set of coefficients $\{a_{ij} : 1 \leq i \leq j \leq n+1\}$, and recall the weight function

$$w(a_{ij}) = t_i^{-1}t_j^{-1}.$$ 

Define a partial order on $\mathcal{K}$ by setting $a_{ij} \leq a_{i'j'}$ if $i \leq i'$ and $j \leq j'$, and on $\mathcal{M}$ by setting $\ell_{ij} \leq \ell_{i'j'}$ if $i \leq i'$ and $j \leq j'$. The significance of this partial order is that if $\alpha, \beta \in \mathcal{K}$ with $\alpha \leq \beta$ and $s \in T'$, then $w(\alpha) \ll w(\beta)$ and similarly $w_L(\alpha) \ll w_L(\beta)$ if $\alpha, \beta \in \mathcal{M}$.

We say that a subset $\mathcal{Z}$ of $\mathcal{K} \cup \mathcal{M}$ is saturated if for any $\alpha \in \mathcal{Z}$, all the $\alpha' \in \mathcal{K} \cup \mathcal{M}$ with $\alpha' \leq \alpha$ are also contained in $\mathcal{Z}$.

Let $\mathcal{D} \subset S(\mathbb{R})$ be a bounded domain such that $B_1 \subset \mathcal{D} \times \mathcal{D}$. We pick positive constants $c_{ij}$ for $1 \leq i \leq j \leq n+1$ and $c'_i$ for $1 \leq i \leq n$ such that:

(a) if $|Yw(a_{ij})| < c_{ij}$, then the $a_{ij}$-coordinate of any integral element in $s(YD)$ is 0;

(b) if $|w_L(\ell_{ij})| < c'_j$, then the $j$th coefficient of $\ell_i$ for any lattice $\Lambda \in \Sigma(L, s)$ is 0;

(c) $c'_i < c'$ for all $i = 1, \ldots, n$, where $c'$ is the constant in Lemma 5.6;

(d) $c_1c_{g+1,n+1} \leq c'_g$, where $c_1$ is the constant in Proposition 5.4;

(e) for any $i \leq i'$ and $j \leq j'$, we have $w(a_{ij})/c_{ij} \leq w(a_{i'j'})/c_{i'j'}$ and $w(\ell_{ij})/c'_j \leq w(\ell_{i'j'})/c'_{j'}$.

More explicitly, we choose $c_{n+1,n+1}$ and $c'_n$ to be sufficiently small and take

$$c_{ij} = \left(\sup_{s \in T'} \frac{w(a_{ij})}{w(a_{n+1,n+1})}\right)c_{n+1,n+1} \quad \text{for } i \leq j \leq n+1;$$

$$c'_i = \left(\sup_{s \in T'} \frac{t_i^{-1}}{t_n^{-1}}\right)c'_n \quad \text{for } i \leq n.$$
For any nondecreasing \( n \)-tuple \( L \) of positive real numbers, and a saturated subset \( Z \) of \( \mathcal{K} \cup \mathcal{M} \), we define the following subset \( T_Z(L, Y) \) of \( T' \):

\[
T_Z(L, Y) := \left\{ s \in T' \mid \begin{array}{ll}
    s_i & \ll X^n \quad \forall i \in \{1, \ldots, n\} \\
    \text{for } a_{ij} \in \mathcal{K}, & |Y w(a_{ij})| < c_{ij} \text{ iff } a_{ij} \in Z \cap \mathcal{K} \\
    \text{for } \ell_{ij} \in \mathcal{M}, & |w_L(\ell_{ij})| < c'_{ij} \text{ iff } \ell_{ij} \in Z \cap \mathcal{M}
\end{array} \right\},
\]

where \( \Theta \) is the absolute constant from Lemma 4.6.

For \( X, Y, L, Z \) as above and any subset \( \mathcal{L} \) of \( W(Z) \), we define the quantity

\[
N(\mathcal{L}, L, Z, X) := \int_{T_Z(L, Y)} \#(A, B) \in (s(YD) \times s(YD)) \cap \mathcal{L} \mid B \in \mathcal{S}(L, s) \delta(s) d^x s.
\]

In the proof of Theorem 5.1, we showed that unless \( Y^{-\Theta_1} < L_1 \) and \( Y^{\Theta_2} > L_n \) for some absolute positive constants \( \Theta_1 \) and \( \Theta_2 \), we have \( \mathcal{S}(L, s) = \emptyset \), which implies that \( N(\mathcal{L}(M), L, Z, X) = 0 \). Therefore,

\[
\mathcal{I}_X(\mathcal{L}(M)) \ll \sum_{\mathcal{L}} \sum_Z N(\mathcal{L}(M), L, Z, X),
\]

where the inner sum is over saturated subsets \( Z \) of \( \mathcal{K} \cup \mathcal{M} \), and the outer sum is over \( n \)-tuples \( L = (L_1, \ldots, L_n) \) with \( L_1 \leq L_2 \leq \cdots \leq L_n \) that partition the region \( \{\mu_1, \ldots, \mu_n\} \in [Y^{-\Theta_1}, Y^{\Theta_2}]^n : \mu_1 \leq \cdots \leq \mu_n \} \) into dyadic ranges.

We may therefore bound the main-body, the shallow-cusp, and the deep-cusp parts of \( \mathcal{I}_X(\mathcal{L}(M)) \) in terms of sums over \( N(\mathcal{L}(M), L, Z, X) \). We have

\[
\begin{align*}
\mathcal{I}_X^{\text{main}}(\mathcal{L}(M)) & \ll \sum_{\mathcal{L}} \sum_{Z: a_{11} \not\in Z} N(\mathcal{L}(M), L, Z, X), \\
\mathcal{I}_X^{\text{scusp}}(\mathcal{L}(M)) & \ll \sum_{\mathcal{L}} \sum_{a_{ij} \in Z} N(\mathcal{L}(M), L, Z, X), \\
\mathcal{I}_X^{\text{dcusp}}(\mathcal{L}(M)) & \ll \sum_{\mathcal{L}} \sum_{Z: a_{g+1, g+1} \in Z} N(\mathcal{L}(M), L, Z, X).
\end{align*}
\]

A preliminary upper bound

We now prove some preliminary results on \( N(\mathcal{L}(1), L, Z, X) \). We start with an upper bound on \( N(\mathcal{L}(1), L, Z, X) \), which also bounds \( N(\mathcal{L}(M), L, Z, X) \) by directly counting the number of possible \( A \)'s and then using the results of §5 to count \( B \)'s. For a saturated subset \( Z \) of \( \mathcal{K} \cup \mathcal{M} \), define

\[
w_L(Z) := \left( \prod_{\alpha \in Z \cap \mathcal{K}} w(\alpha) \right) \left( \prod_{\alpha \in Z \cap \mathcal{M}} w_L(\alpha) \right).
\]

In what follows, the \( n \)-tuple \( L \) will be clear from the context, and we simply write \( w \) in place of \( w_L \).

**Proposition 6.3** Suppose that \( Z \) is a saturated subset of \( \mathcal{K} \cup \mathcal{M} \). Then

\[
N(\mathcal{L}(1), L, Z, X) \ll X^{n+1} \int_{T_Z(L, Y)} Y^{-\#(Z \cap \mathcal{K})} w(Z)^{-1} \left( \prod_{1 \leq i < j \leq n} \frac{L_i}{L_j} \right) \delta(s) d^x s.
\]

27
To this end, define the set $\mathcal{Z}$. Suppose now that Condition (b) is satisfied. Then the determinants of the matrices obtained by replacing the $i$-th column (and row) of $B$ for any $0 \leq i < n$ by translating the vector $v$ is singular. The kernel of $B$ has dimension at least 2.

Proof: By Proposition 4.4, the number of elements $A \in s(\mathcal{Y} \mathcal{D}) \cap S(\mathcal{Z})$ is

$$\ll Y^{(n+1)(n+2)/2} \prod_{a_{ij} \in \mathcal{Z} \cap \mathcal{K}} (Yw(a_{ij}))^{-1} \ll Y^{(n+1)(n+2)/2-\#(\mathcal{Z} \cap \mathcal{K})}w(\mathcal{Z} \cap \mathcal{K})^{-1}. \quad (41)$$

By the definition of $T_\mathcal{Z}(L, Y)$, it follows from (28) that for every $s \in T_\mathcal{Z}(L, Y)$, we have

$$\# \Sigma(L, s) \ll (L_1 \cdots L_n)^{n+1} w(\mathcal{Z} \cap \mathcal{M})^{-1} \prod_{1 \leq i < j \leq n} \frac{L_i}{L_j}. \quad (42)$$

For each $\Lambda \in \Sigma(L, s)$, Proposition 5.5 implies that the number of integral symmetric matrices $B \in s(\mathcal{Y} \mathcal{D})$ whose row space is contained in $\Lambda$ is

$$\ll \frac{Y^{n(n+1)/2}}{(L_1 \cdots L_n)^{n+1}} \prod_{1 \leq i < j \leq n+1} \frac{L_j}{L_i}. \quad (43)$$

Combining (41), (42), and (43), and recalling that $X = Y^{n+1}$, gives (40). □

Conditions for vanishing discriminant

Next, we give some conditions on $\mathcal{Z}$ that ensure $N(\mathcal{L}(1), L, \mathcal{Z}, X) = 0$. We start with the following algebraic result that gives sufficient conditions on a pair $(A, B) \in W(\mathcal{C})$ that ensure it has discriminant 0.

Lemma 6.4 Suppose that $(A, B)$ is an element of $W(\mathcal{C})$ such that one of the following three conditions are satisfied:

(a) The kernel of $B$ has dimension at least 2.

(b) There is a nonzero vector $v \in \mathbb{C}^{n+1}$ that is in the kernel of $B$ and isotropic with respect to $A$.

(c) There exists $k \in \{1, \ldots, g+1\}$ such that $a_{ij} = b_{ij} = 0$ for all $1 \leq i \leq k$ and all $1 \leq j \leq n+1-k$.

Then $\Delta(A, B) = 0$.

Proof: This is a standard result in the algebraic geometric theory of pencils of quadrics. We give another proof using the explicit formula for $f(x, y) = f_{A, B}(x, y)$. The claim regarding Condition (c) is Lemma 4.5. If the kernel of $B$ has dimension at least 2, then the quadratic form defined by $A$ restricted to the kernel of $B$ admits a nonzero isotropic vector in $\mathbb{C}^{n+1}$. Thus Condition (a) implies Condition (b). Suppose now that Condition (b) is satisfied. Then the $y^{n+1}$-coefficient of $f(x, y)$ is 0 since $B$ is singular. The $xy^n$-coefficient of $f(x, y)$ equals, up to sign, the alternating sum of the determinants of the matrices obtained by replacing the $i$-th column of $B$ by the $i$-th column of $A$. By translating the vector $v$ to $(1, 0, 0, \ldots, 0)$ using an element of $\text{SL}_{n+1}(\mathbb{C})$, we may assume that the first column (and row) of $B$ is 0 and the $(1, 1)$-entry of $A$ is 0. It is then easy to see that the determinant of the matrix obtained by replacing the $i$-th column of $B$ by the $i$-th column of $A$ is 0 for any $i$. Hence $\Delta(A, B) = \Delta(f) = 0$. □

We now translate these conditions into the vanishing of $N(\mathcal{L}(1), L, \mathcal{Z}, X)$ for certain sets $\mathcal{Z}$. To this end, define the set $\mathcal{Z}_1 \subset \mathcal{K} \cup \mathcal{M}$ by

$$\mathcal{Z}_1 := \{a_{ij} \mid i \leq j, i + j \leq n\} \cup \{\ell_{ij} \mid i + j \leq n + 1\}.$$
Lemma 6.5 Let \( Z \) be a saturated subset of \( K \cup \mathcal{M} \) satisfying one of the following two conditions:

(a) The set \( Z \) is not contained in \( Z_1 \).

(b) There exists \( k \in \{1, \ldots, g+1\} \) such that \( a_{kk} \in \mathbb{Z} \) and \( \ell_{n+1-k,k} \in \mathbb{Z} \).

Then \( N(\mathcal{L}(1), L, \mathbb{Z}, X) = 0 \).

Proof: If \( Z \) contains some \( \ell_{ij} \notin Z_1 \), then for every \( s \in T_Z(L, Y) \), the set \( \Sigma(L, s) \) (and hence \( S(L, s) \)) is empty by Lemma 5.6. This implies that \( N(\mathcal{L}(1), L, \mathbb{Z}, X) = 0 \). If \( Z \) contains some \( a_{ij} \notin Z_1 \), then every integral \((A, B) \in s(YD) \times s(YD)\) has discriminant \( 0 \) by Condition (c) of Lemma 6.4. Once again, this implies that \( N(\mathcal{L}(1), L, \mathbb{Z}, X) = 0 \).

Let \( k \) be an integer satisfying Condition (b) of the lemma, and let \( s \in T_Z(L, Y) \). Let \((A, B)\) be such that \( A \in s(YD) \) and \( B \in S(L, s) \). Since \( \ell_{n+1-k,k} \in \mathbb{Z} \), it follows that there exists a nonzero vector \( v \in \mathbb{C}^{n+1} \) of the form \((v_1, \ldots, v_k, 0, \ldots, 0)\) that is in the kernel of \( B \). Since \( a_{kk} \in \mathbb{Z} \), it follows that \( v \) is isotropic with respect to \( A \). By Condition (b) of Lemma 6.4, it follows that \( \Delta(A, B) = 0 \), implying that \( N(\mathcal{L}(1), L, \mathbb{Z}, X) = 0 \), as desired. \(\square\)

6.2 Bounding the number of distinguished elements in the main body

In this subsection, we bound the number of distinguished elements in the main body:

Theorem 6.6 We have \( I^\text{main}_X(\mathcal{L}(1)) = O(X^{n+1-1/(10n)}) \).

As \( \mathcal{L}(M) \subset \mathcal{L}(1) \) for \( M \geq 1 \), it follows that \( I^\text{main}_X(\mathcal{L}(M)) \) satisfies the same bound.

We will use the Selberg sieve to show that distinguished elements are negligible in number in the main body. However, applying the Selberg sieve requires asymptotics along with a power saving error term. Our methods in §5 do not yield such results.

Hence we will instead fiber over \( B \in s(YD) \cap S(\mathbb{Z}) \) having determinant \( 0 \), apply the Selberg sieve to prove that there are negligibly many \( A \in s(YD) \cap S(\mathbb{Z}) \) such that \((A, B)\) is distinguished, and then bound the number of possible \( B \)'s using the results of Section 5. To carry out the middle step, we require the following lower bound on the number of nondistinguished elements modulo primes \( p \) that is independent of \( p \) and \( B \).

Lemma 6.7 Let \( B_0 \) be an element in \( S(\mathbb{F}_p) \) with \( \mathbb{F}_p \)-rank \( n \). Let \( \mathcal{S}^\text{ndist}(\mathbb{F}_p) \) denote the set of elements \( A \in S(\mathbb{F}_p) \) such that \((A, B)\) has nonzero discriminant and \( A \) and \( B \) do not have a common isotropic \((g+1)\)-dimensional subspace. Then

\[
\frac{\# \mathcal{S}^\text{ndist}(\mathbb{F}_p)}{\# S(\mathbb{F}_p)} \gg n 1.
\]

Proof: For an element \( B \in S(\mathbb{F}_p) \) with \( \mathbb{F}_p \)-rank \( n \) and kernel spanned by \( v \), let \( d(B) \) denote the discriminant of the corresponding quadratic form on \( \mathbb{F}_p^{n+1}/(\mathbb{F}_p v) \). If \( B_1, B_2 \in S(\mathbb{F}_p) \) have \( \mathbb{F}_p \)-rank \( n \) and \( d(B_1)/d(B_2) \in \mathbb{F}_p^2 \), then \( B_1 \) and \( B_2 \) are \( \text{SL}_{n+1}(\mathbb{F}_p) \)-equivalent. Indeed, by using \( \text{SL}_{n+1}(\mathbb{F}_p) \) transformations, we may assume the last row and columns of \( B_1 \) and \( B_2 \) are all 0. The nondegenerate forms defined by the top left \( n \times n \) blocks of \( B_1 \) and \( B_2 \) have discriminants \( d(B_1) \) and \( d(B_2) \), which are in the same quadratic residue class. Hence they are equivalent via an element \( \gamma \in \text{GL}_n(\mathbb{F}_p) \). Expanding \( \gamma \) to an element in \( \text{SL}_{n+1}(\mathbb{F}_p) \) by appending an additional row and column whose entries are all 0, except for the \((n+1, n+1)\)-entry which is \( \det \gamma^{-1} \), gives an element in \( \text{SL}_{n+1}(\mathbb{F}_p) \) that takes \( B_1 \) to \( B_2 \).
Let $B_0 \in S(F_p)$ have $\mathbb{F}_p$-rank $n$. For each binary $n$-ic form $f(x, y) = a_0x^n + \cdots + a_ny^n$ over $\mathbb{F}_p$ that splits completely over $\mathbb{F}_p$ such that $\Delta(xf(x, y)) \neq 0$ and $a_0 \neq 0$, we construct a nondistinguished element $(A_0, B_0)$ with $f_{A_0B_0} = xf(x, y)$. Let $f$ be such a form. Then $a_n \neq 0$. Let $\alpha = d(B_0)/a_n$. As noted in §3.1, there exist at least two (in fact $2^{n-1}$) $\text{SL}_n(\mathbb{F}_p)$-orbits of $(A, B) \in W_n(\mathbb{F}_p)$ such that $f\gamma A, B = \alpha f(x, y)$. Pick two inequivalent representatives $(A_1, B_1)$ and $(A_2, B_2)$. Let $A'_1$ and $A'_2$ be the $(n+1)$-ary quadratic forms obtained from $A_1$ and $A_2$, respectively, by appending an additional row and column whose entries are all zero except for the $(n+1, n+1)$-entry which is $-1$. Let $B'_1$ and $B'_2$ be the $(n+1)$-ary quadratic forms obtained from $B_1$ and $B_2$, respectively, by appending an additional row and column whose entries are all zero except for the $(n+1, n+1)$-entry which is $-1$. Then $f_{A'_1, B'_1} = f_{A'_2, B'_2} = xf(x, y)$. Since $(A_1, B_1)$ and $(A_2, B_2)$ are $\text{SL}_n(\mathbb{F}_p)$-inequivalent, it follows that $(A'_1, B'_1)$ and $(A'_2, B'_2)$ are $\text{SL}_{n+1}(\mathbb{F}_p)$-inequivalent. Hence, without loss of generality, we may assume that $(A'_1, B'_1)$ is nondistinguished. Now $d(B'_1) = a_n = d(B_0)$, and so there exists $\gamma \in \text{SL}_{n+1}(\mathbb{F}_p)$ such that $\gamma B'_1 \gamma^t = B_0$. Then $A_0 = \gamma A'_1 \gamma^t$ does the job.

We complete the proof of the lemma via the orbit-stabilizer theorem. By the above construction, there are $\gg_n p^{n+1}$ binary $(n+1)$-ic forms $xf(x, y)$, with $\Delta(xf(x, y)) \neq 0$ and $a_0 \neq 0$, such that there exists an element $A \in S(\mathbb{F}_p)$ with $f_{A, B_0} = xf(x, y)$ and $(A, B_0)$ nondistinguished. The group $G_{B_0}(\mathbb{F}_p) = \{ \gamma \in \text{SL}_{n+1}(\mathbb{F}_p) : \gamma B_0 \gamma^t = B_0 \}$ acts on the set of such $A$ with stabilizer of size $\#J_{xf}[2](\mathbb{F}_p)$, where $J_{xf}$ is the Jacobian of the hyperelliptic curve defined by $z^2 = xf(x, y)y$. Any element of $\gamma \in G_{B_0}(\mathbb{F}_p)$ preserves the kernel $\mathbb{F}_p v$ of $B_0$ and stabilizes the nondegenerate form $b_0$ on $\mathbb{F}_p^{n+1}/(\mathbb{F}_p v)$ induced by $B_0$. The determinant 1 condition then gives

$$\#G_{B_0}(\mathbb{F}_p) = \#O(b_0)(\mathbb{F}_p) = 2p^{\frac{n^2+2n}{2}} \left(1 - \frac{O(1)}{p^2}\right).$$

Finally, since $\#J_{xf}[2](\mathbb{F}_p) \ll_n 1$, we have

$$\#S_{B_0}^{\text{ndist}}(\mathbb{F}_p) \gg_n p^{n+1} p^{a(n+1)/2} = p^{(n+1)(n+2)/2} = \#S(\mathbb{F}_p),$$

as desired. \(\square\)

**Corollary 6.8** Fix $a \in \mathbb{F}_p^\times$ and $B_0 \in S(\mathbb{F}_p)$ with rank $n$. Let $S_{B_0}^{\text{ndist}}(\mathbb{F}_p)_{a_{11} = a}$ denote the set of all elements $A \in S_{B_0}^{\text{ndist}}(\mathbb{F}_p)$ with $a_{11} = a$. Then

$$\frac{\#S_{B_0}^{\text{ndist}}(\mathbb{F}_p)_{a_{11} = a}}{\#S(\mathbb{F}_p)/p} \gg_n 1.$$

**Proof:** Since the property of $(A, B)$ being nondistinguished is preserved when $A$ is multiplied by an element of $\mathbb{F}_p^\times$, the claim follows immediately from Lemma 6.7. \(\square\)

We now bound the number of pairs $(A, B)$ in the main body where the first row and column of $B$ are zero.

**Proposition 6.9** We have

$$\int_{s \in \mathcal{T}'} \#\{ (A, B) \in (s(Y^D) \times s(Y^D)) \cap L(1) : b_{ii} = 0 \forall i \} \delta(s)d^x s \ll X^{n+1-1/(10n)}.$$  \(44\)

**Proof:** Let $s \in \mathcal{T}'$ be an element with $Yw(a_{11}) \gg 1$. Then

$$\#\{ A \in s(Y^D) \cap S(\mathbb{Z}) \} \ll Y^{(n+1)(n+2)/2};$$

$$\#\{ B \in s(Y^D) \cap S(\mathbb{Z}) : b_{ii} = 0 \forall i \} \ll Y^{n(n+1)/2} \prod_{i=1}^{n+1} w(b_{ii})^{-1}.$$

30
For each $B \in s(YD) \cap S(Z)$ having rank $n$, we bound the number of $A \in s(YD) \cap S(Z)$ such that $(A, B)$ is distinguished. Indeed, after additionally fibering over the coefficient $a_{11}$, Corollary 6.6 in conjunction with an application of the Selberg sieve, used as in [28], saves a 1/5th power of the smallest coefficient range among the remaining $a_{ij}$. That is, we obtain a saving of $(Yw(a_{12}))^{-1/5}$.

Therefore, the left hand side of (44) is

\[
\ll Y^{(n+1)^2-1/5} \int_{s \in \mathcal{T}} w(a_{12})^{-1/5} s_1^{n(n+1)} s_2^{(n-1)(n+1)} \cdots s_n^{n+1} \delta(s) d^x s
\ll Y^{(n+1)^2-1/5} \int_{s \in \mathcal{T}} s_1^{(n-1)/5} \prod_{j=2}^n s_j^{(2n+2-2j)/5-(n+1)(n+1-j)(j-1)} d^x s.
\]

In particular, the power of $s_i$ above is negative for all $i \in \{2, \ldots, n\}$, and hence the integral over $s_2, \ldots, s_n$ is absolutely bounded. The condition that $Yw(a_{11}) \gg 1$ on the integrand implies that we have $s_1 \ll Y^{1/2n}$. Therefore the terms in (45) are

\[
\ll Y^{(n+1)^2-1/5} \int_{s_1 \ll s_1 \ll Y^{1/(2n)}} s_1^{(n-1)/5} d^x s_1 \ll Y^{(n+1)^2-1/5+(n-1)/(10n)} = Y^{(n+1)^2-(n+1)/(10n)}.
\]

Since $Y = X^{1/(n+1)}$, we obtain the result. □

**Remark 6.10** Our use of the Selberg sieve saves a power of the smallest range of any coordinate. In the above proof, we fiber over $a_{11}$ because in the region of the main body close to the cusp, just before we enter the shallow cusp, the range of $a_{11}$ has size $\ll 1$. In this case, the Selberg sieve gives no saving at all. Once we fiber over $a_{11}$, the next smallest range is that of $a_{12}$. Implicit in our proof is an argument that either the range of $a_{12}$ is large, in which case the Selberg sieve gives the desired saving, or the number of pairs $(A, B)$ is automatically small.

**Proof of Theorem 6.6:** Recall from (39) that we have

\[
T_X^{\text{main}}(\mathcal{L}(1)) \ll \sum_{L} \sum_{Z, a_{11} \notin Z} N(\mathcal{L}(1), L, Z, X),
\]

where the second sum is over all saturated $Z$. Since $Z$ is saturated and $a_{11} \notin Z$, we have $Z \subset M$. If $\ell_{k,1} = 0$ for every $k = 1, \ldots, n$, then $(1, 0, \ldots, 0)$ is in the kernel of $B$ implying that the top row of $B$ is zero. The number of such pairs $(A, B)$ has already been bounded in Proposition 6.9, and hence we may assume that $\ell_{n,1} \notin Z$. Fix a nondecreasing $n$-tuple $L$ of positive real numbers, and a saturated $Z \subset M$ with $\ell_{n,1} \notin Z$ such that $N(\mathcal{L}(1), L, Z, X) \neq 0$. We partition the integrand $T_Z(L, Y)$ into two parts: let $T_1$ denote the subset of $T_Z(L, Y)$ consisting of elements $s$ for $s = (s_i)_i$ with $s_n \geq Y^{\delta}$, and let $T_2$ denote the subset of elements $s$ with $1 \ll s_n < Y^{\delta}$, where $\delta$ is a positive constant to be optimized later.

We first bound the contribution to $N(\mathcal{L}(1), L, Z, X)$ from $T_1$. Since $Yw(a_{11}) \gg 1$, we have

\[
\#((s(YD) \times s(YD)) \cap \mathcal{L}(1)) \leq \#((s(YD) \times s(YD)) \cap W_{n+1}(Z)) \ll Y^{(n+1)(n+2)}
\]

for $s \in T_1$. Integrating over $T_1$ gives the bound

\[
\int_{s \in T_1} \#((s(YD) \times s(YD)) \cap \mathcal{L}(1)) \delta(s) d^x s \ll Y^{(n+1)(n+2)} \int_{s_1, \ldots, s_{n-1} \gg 1} \int_{s_n \geq Y^{\delta}} \delta(s) d^x s
\ll Y^{(n+1)(n+2)} \int_{s_n \geq Y^{\delta}} s_n^{n(n+1)} d^x s_n
\ll Y^{n+1-n(n+1)\delta} X^{n+1}.
\]
Next, we consider the contribution from $T_2$. Define the map $\pi : Z_1 \cap M \to M$ by

$$\pi(\ell_{ij}) = \begin{cases} 
\ell_{n1} & \text{if } j = 1 \text{ and } i \geq 2, \\
\ell_{i,n+2-i} & \text{otherwise}.
\end{cases}$$

Since we have assumed that $N(L(1), L, Z, X) \neq 0$, Lemma 6.5 implies that $Z \subset Z_1$ and so the image of $\pi$ lies in $M \setminus Z$. Then for any $\alpha \in Z_1 \cap M$ and any $s \in T_2(L, Y)$, we have $w_L(\pi(\alpha)) \gg w_L(\alpha)$ and $w_L(\pi(\alpha)) \gg 1$. These inequalities along with (42) and (43) imply that for any $s \in T_2(L, Y)$, the number $\#((s(YD) \cap S(L, s))$ of possible $B'$s is

$$\ll Y^{n(n+1)/2}w(Z \cap M)^{-1}\left( \prod_{1 \leq i, j \leq n; i+j > n+1} \frac{L_i}{L_j} \right) \ll Y^{n(n+1)/2}w(a_{11})^{-1/5}w(\pi(\ell_{ii}))^{-1/5}\left( \prod_{i \in Z_1 \cap M; \ell \neq \ell_{n1}} \frac{w(\pi(\ell))}{w(\ell)}\right)\left( \prod_{1 \leq i, j \leq n; i+j > n+1} \frac{L_i}{L_j} \right).$$

For each possible $B$, applying the Selberg sieve using Lemma 6.7 gives us a bound of

$$\ll Y^{(n+1)(n+2)/2}Y^{-1/5}w(a_{11})^{-1/5}$$

for the number of possible choices for $A$. Therefore,

$$\#((s(YD) \times s(YD)) \cap L(1)) \ll Y^{-1/5}X^{n+1}w(a_{11})^{-1/5}\left( \prod_{i \in Z_1 \cap M; \ell \neq \ell_{n1}} \frac{w(\pi(\ell))}{w(\ell)}\right)\left( \prod_{1 \leq i, j \leq n; i+j > n+1} \frac{L_i}{L_j} \right),$$

for $s \in T_2$. We compute the ratio of these weights: For any $i \geq 2$ and $j = 1$, we have

$$\frac{w(\pi(\ell_{ii}))}{w(\ell_{ii})} = \frac{w(\ell_{n1})}{w(\ell_{i1})} = \frac{L_n}{L_i}.$$

For any other $i, j$, we have

$$\frac{w(\pi(\ell_{ij}))}{w(\ell_{ij})} = \frac{w(\ell_{i,n+2-i})}{w(\ell_{ij})} = \frac{t_j}{t_{n+2-i}}.$$

As the $L_i$ are nondecreasing and positive, we multiply by the Haar measure character $\delta(s)$ to obtain

$$w(a_{11})^{-1/5}\left( \prod_{i \in Z_1 \cap M; \ell \neq \ell_{n1}} \frac{w(\pi(\ell))}{w(\ell)}\right)\left( \prod_{1 \leq i, j \leq n; i+j > n+1} \frac{L_i}{L_j} \right)\delta(s) \leq w(a_{11})^{-1/5}\left( \prod_{i, j \geq 1; i+j \leq n+1} \frac{t_j}{t_{n+2-i}}\right)\left( \prod_{i = 2}^n \frac{t_{n+2-i}}{t_i}\right)\left( \prod_{1 \leq j < i \leq n+1} \frac{t_i}{t_j} \right) \leq w(a_{11})^{-1/5}\left( \prod_{i = 2}^n \frac{t_{n+2-i}}{t_i}\right) \leq w(a_{11})^{-1/5}\frac{1}{s_1^{(n+1)(n-1)}s_2^{-(n+1)(n-2)} \cdots s_{n-1}^{-(n+1)}}.$$

The powers of $s_i$ in the above expression are negative for $1 \leq i \leq n-1$, while the power of $s_n$ is $2/5$. Integrating over $T_2$ now gives the bound

$$\int_{s \in T_1} \#((s(YD) \times s(YD)) \cap L(1)) \delta(s)d^x s \ll Y^{-1/5}X^{n+1}\int_{1 \leq s_n \ll Y^\delta} s_n^{2/5} d^x s_n \ll Y^{-1/5+(2\delta)/5}X^{n+1}.$$ (47)
Combining (46) and (47) and choosing \( \delta = \frac{5n+6}{5n^2+5n+2} \) yields
\[
N(\mathcal{L}(1), L, Z, X) \ll X^{n+1-\frac{n-2}{5n^2+5n+2}}.
\]
The summation of this bound over the \( O(1) \) different possible \( L \)'s and the \( O(Y^\epsilon) \) different possible \( L \)'s, in conjunction with the bound in Proposition 6.9, implies Theorem 6.6. \( \square \)

6.3 Bounding the number of distinguished elements in the shallow cusp

In this subsection, we bound the number of distinguished elements having large \( q \)-invariant that lie in the shallow cusp of the fundamental domain.

**Theorem 6.11** Let \( \eta > 0 \) be any real number. Assume that \( M > X^\eta \). Then
\[
I_X^{\text{scusp}}(\mathcal{L}(M)) = O(X^{n+1-\min(\eta,1)/(2n^6)}).
\]

We will take \( \eta = 1/4 \) when we prove Theorem 5 in §6.5.

6.3.1 A preliminary bound of \( O_\epsilon(X^{n+1+\epsilon}) \)

We again use (39) to write
\[
I_X^{\text{scusp}}(\mathcal{L}(M)) \ll \sum_{L,Z} N(\mathcal{L}(M), L, Z, X),
\]
where the sum is over nondecreasing \( n \)-tuples \( L = (L_1, \ldots, L_n) \) of positive real numbers that partition the region \( \{(\mu_1, \ldots, \mu_n) \in [Y^{-\Theta_1}, Y^\Theta_2]^n : \mu_1 \leq \mu_2 \leq \ldots \leq \mu_n \} \) into dyadic ranges, and over saturated \( Z \subset K \cup M \) such that \( a_{11} \in Z \) and \( a_{g+1,g+1} \notin Z \). By Lemma 6.5, we have \( N(\mathcal{L}(M), L, Z, X) > 0 \) only when \( Z \subset Z_1 \), which we henceforth assume.

For \( k \in \{0, \ldots, g\} \), define the map \( \pi_k : Z_1 \to K \cup M \) by
\[
\pi_k(a_{ij}) = a_{n+1-j,j}, \quad \pi_k(\ell_{ij}) = \begin{cases} 
\ell_{n+1-j,j} & \text{if } i > j \text{ and } j \leq k, \\
\ell_{i,n+2-i} & \text{otherwise}.
\end{cases}
\]

We define the auxiliary set \( Z^* \) by
\[
Z^* = \{a_{ij} \mid i \leq j, i + j \leq n\} \cup \{\ell_{ij} \mid i \leq j, i + j \leq n + 1\} = Z_1 \setminus \{\ell_{ij} \mid i > j, i + j \leq n + 1\}.
\]

Then, when restricted to \( Z^* \subset Z_1 \), the functions \( \pi_k \) are equal for every \( k \).

**Lemma 6.12** For any \( k \in \{0, \ldots, g\} \), we have
\[
\left( \prod_{\alpha \in Z^*} \frac{w(\pi_k(\alpha))}{w(\alpha)} \right) \delta(s) = 1. \tag{48}
\]

**Proof:** We directly compute
\[
\prod_{\alpha \in Z^*} \frac{w(\pi_k(\alpha))}{w(\alpha)} = \left( \prod_{i \leq j \leq n+1} \frac{w(a_{n+1-j,j})}{w(a_{ij})} \right) \left( \prod_{i \leq j \leq n+2} \frac{w(\ell_{i,n+2-i})}{w(\ell_{ij})} \right)
\]
\[
= \left( \prod_{i \leq j \leq n+1} \frac{t_i}{t_{n+1-j}} \right) \left( \prod_{i \leq j \leq n+2} \frac{t_j}{t_{n+2-i}} \right)
\]
\[
= \left( \prod_{1 \leq r} \frac{t_i}{t_r} \right) \left( \prod_{j < r \leq n+2} \frac{t_j}{t_r} \right),
\]

33
which is $\delta(s)^{-1}$. □

Fix a saturated set $\mathcal{Z} \subset \mathcal{Z}_1$ such that $a_{11} \in \mathcal{Z}$, $a_{g+1, g+1} \notin \mathcal{Z}$ and $N(\mathcal{L}(M), L, Z, X) > 0$. Let $k \in \{1, \ldots, g\}$ be the largest integer such that $a_{kk} \in \mathcal{Z}$. Then we have the following results.

**Lemma 6.13** Let $\mathcal{Z}$ and $k$ be as above. Then for every $\alpha \in \mathcal{Z}$, we have $\pi_k(\alpha) \notin \mathcal{Z}$. In particular, for any $s \in T_Z(L, Y)$, we have $Y w(\pi_k(\alpha)) \gg 1$.

**Proof:** Since $a_{n+1-j, j} \notin \mathcal{Z}$ for any $j$ and $\mathcal{Z} \subset \mathcal{Z}_1$, we have $\pi_k(\alpha) \notin \mathcal{Z}$ for any $a_{ij} \in \mathcal{Z}$. Moreover, since $a_{jj} \in \mathcal{Z}$ for every $j \leq k$, it follows from Lemma 6.5 that $\ell_{n+1-j, j} \notin \mathcal{Z}$. Furthermore, $\ell_{n+2-i} \notin \mathcal{Z}_1$. Hence $\pi_k(\ell_{ij}) \notin \mathcal{Z}$ for any $\ell_{ij} \in \mathcal{Z}$. □

**Lemma 6.14** Let $\mathcal{Z}$ and $k$ be as above. Then, uniformly for $s \in T_Z(L, Y)$, we have

$$
\left(\prod_{\alpha \in \mathcal{Z}} \frac{w(\pi_k(\alpha))}{w(\alpha)}\right) \left(\prod_{1 \leq i < j \leq n} \frac{L_i}{L_j}\right) \delta(s) \ll 1. \tag{49}
$$

**Proof:** Since we have

$$
\frac{w(a_{n+1-j, j})}{w(a_{ij})} = \frac{t_i}{t_{n+1-j}}, \quad \frac{w(\ell_{i,n+2-i})}{w(\ell_{ij})} = \frac{t_j}{t_{n+2-i}}, \quad \frac{w(\ell_{n+1-j, j})}{w(\ell_{ij})} = \frac{L_{n+1-j}}{L_i},
$$

it follows that $w(\pi_k(\alpha))/w(\alpha) \gg 1$ for every $k$, $\alpha \in \mathcal{Z}_1$, and $s \in T_Z(L, Y)$. Thus, by adding elements in $Z_1$ to $Z$, if necessary, we can assume that $\mathcal{Z}$ is equal to

$$
\mathcal{Z}_0 = \{a_{ij} : i \leq j, i \leq k, i+j \leq n\} \cup \{\ell_{ij} : i > j > k, i+j \leq n+1\}
\cup \{\ell_{ij} : i > j, j \leq k, i+j \leq n+1\} \cup \{\ell_{ij} : i \leq j, i+j \leq n+1\}.
$$

Denote the four sets on the right hand side of the above equation as $S_1$, $S_2$, $S_3$, and $S_4$, respectively. For an element $\ell_{ij} \in S_2$, we have

$$
\frac{w(\pi_k(\ell_{ij}))}{w(\ell_{ij})} = \frac{w(\ell_{i,n+2-i})}{w(\ell_{ij})} = \frac{t_j}{t_{n+2-i}} = \frac{w(\pi_k(a_{j,i-1}))}{w(a_{j,i-1})}.
$$

Therefore,

$$
\left(\prod_{\alpha \in \mathcal{Z}_0} \frac{w(\pi_k(\alpha))}{w(\alpha)}\right) \left(\prod_{1 \leq i < j \leq n} \frac{L_i}{L_j}\right) \delta(s) = \left(\prod_{\alpha \in \mathcal{Z}_1} \frac{w(\pi_k(\alpha))}{w(\alpha)}\right) \left(\prod_{\alpha \in S_3} \frac{w(\pi_k(\alpha))}{w(\alpha)}\right) \left(\prod_{1 \leq i < j \leq n} \frac{L_i}{L_j}\right) \delta(s)
= \left(\prod_{i \geq j \leq k} \frac{L_{n+1-j}}{L_i}\right) \left(\prod_{1 \leq i < j \leq n} \frac{L_i}{L_j}\right)
= \prod_{1 \leq i < j \leq n} \frac{L_i}{L_j}
\leq 1,
$$

where the second equality follows from Lemma 6.12, and the last inequality follows because the $L_i$’s are nondecreasing. □

Proposition 6.3 and Lemmas 6.13 and 6.14 thus yield the bound

$$
N(\mathcal{L}(M), L, Z, X) \ll X^{n+1} \int_{1 \ll s_1, \ldots, s_n \ll X^{\Theta}} d^x s \ll_{\epsilon} X^{n+1+\epsilon}.
$$

We now work towards obtaining a power saving.
6.3.2 Strategy towards a power saving

In light of Proposition 6.3, it is enough to have a bound of the form

\[ Y^{-\#Z} w(Z)^{-1} \left( \prod_{1 \leq i < j \leq n} \frac{L_i}{L_j} \right) \delta(s) \ll X^{-\delta} \]  

(50)

for some \( \delta > 0 \), for all \( s \in T_Z(L,Y) \). By modifying \( \pi_k \) on a certain subset of \( Z \), we are able to obtain (50) except for some \( s \in T_Z(L,Y) \) satisfying some special conditions. We then consider the contribution from these special \( s \) using a different count for \( \#((s(Y)D) \times s(Y)D) \cap L(M) \).

More precisely, let \( K_1 := \{a_{ij} : 1 \leq j \leq g+2 \} \). Then \( K_1 \) consists exactly of those \( \alpha \in K \) such that the exponent of every \( s_i \) is negative in \( w(\alpha) \). As such, one expects that the hardest case is when \( Z = K_1 \). We show first in Lemma 6.15 how to reduce to considering only \( Z \cap K_1 \).

**Lemma 6.15** Let \( Z \subset Z_1 \) be saturated with \( a_{11} \in Z, a_{g+1,g+1} \notin Z \) and \( N(L(M),L,Z,X) > 0 \). For any \( Z' \subset K_1 \) and any \( s \in T' \), we write

\[ I(Z',s) = Y^{-\#Z'} w(Z')^{-1} \prod_{i=1}^{g} s_i^{-((n+1)(g+2)} \prod_{i=g+1}^{n-1} s_i^{-(n+1)(n-i)} \]

Then for any \( s \in T_Z(L,Y) \), we have

\[ Y^{-\#Z} w(Z)^{-1} \left( \prod_{1 \leq i < j \leq n} \frac{L_i}{L_j} \right) \delta(s) \ll I(Z \cap K_1, s) \]

We then prove in Lemma 6.16 the following bound for \( I(Z \cap K_1) \) when \( Z \cap K_1 \) is a proper subset of \( K_1 \), which gives a bound of the form (50) when \( s_n \ll Y^{1/2-\delta} \).

**Lemma 6.16** Let \( Z \subset Z_1 \) be saturated with \( a_{11} \in Z, a_{g+1,g+1} \notin Z \) and \( N(L(M),L,Z,X) > 0 \). Suppose \( Z \cap K_1 \neq K_1 \). For any \( s \in T_Z(L,Y) \), if \( I(Z \cap K_1, s) \gg Y^{-2\delta} \), then \( s_n \gg Y^{1/2-\delta} \).

In the case where \( s_n \gg Y^{1/2-\delta} \), the Haar measure turns out to be very small so we may simply ignore the singularity condition of \( B \) and prove the following bound.

**Lemma 6.17** Let \( Z \subset Z_1 \) be saturated with \( a_{11} \in Z, a_{g+1,g+1} \notin Z \) and \( N(L(M),L,Z,X) > 0 \). Suppose \( Z \cap K_1 \neq K_1 \). Then for any \( s \in T_Z(L,Y) \) with \( s_n \gg Y^{1/2-\delta} \),

\[ \#((s(Y)D) \times s(Y)D) \cap W(Z)) \delta(s) \ll Y^{(\frac{g+2}{2} + n(n+1)+n(2n^2+9n+9))} \delta \]

Therefore, by taking \( \delta = (n-2)/(4n^2+14n+4) \), we obtain the following result from Proposition 6.3 and Lemmas 6.15, 6.16, and 6.17:

**Proposition 6.18** Let \( Z \subset Z_1 \) be saturated with \( a_{11} \in Z, a_{g+1,g+1} \notin Z \) and \( N(L(M),L,Z,X) > 0 \). Suppose \( Z \cap K_1 \neq K_1 \). Then

\[ N(L(M),L,Z,X) \ll X^{n+1-\frac{n^2-2}{2(n+1)(n^2+7n+7)}} \]

We next handle the case \( K_1 \subset Z \). We give necessary conditions in Lemma 6.19 on \( s \) so that a bound of the form (50) does not hold.
Lemma 6.19 Let \( \mathcal{Z} \subset \mathcal{Z}_1 \) be saturated with \( a_{11} \in \mathcal{Z}, a_{g+1,g+1} \notin \mathcal{Z} \) and \( N(\mathcal{L}(M), L, \mathcal{Z}, X) > 0 \). Suppose \( K_1 \subset \mathcal{Z} \). For any \( s \in T_\mathcal{Z}(L,Y) \), if \( I(K_1, s) \gg X^{-\delta} \), then
\[
Y^{-\delta} \ll \frac{s_i}{s_{n-i}} \ll Y^\delta, \quad \text{for} \quad i = 1, \ldots, g-1
\]
\[
Y^{1/2-(g/2)\delta} R^{-1} \ll \frac{s_g}{s_{g+1}} \ll Y^{1/2+3g\delta} R^{-1}
\]
\[
1 \ll s_{g+1} \ll Y^\delta
\]
\[
Y^{1/2-\delta} R^{-1} \ll s_{g+2} \ll Y^{1/2+g\delta} R^{-1}
\]
where
\[
R = \prod_{i=g+3}^n s_i \ll Y^{1/2+3g\delta}.
\]
Note that the coefficients of \( \delta \) in the exponents in the above bounds are not optimal and are simply chosen to make the formula look nice. The optimal coefficients can be obtained from the proof.

When \( s \) satisfies (51), we give further conditions in Lemma 6.20 on \( s \) so that simply using the Haar measure and ignoring the singularity condition by counting all symmetric matrices is not enough for a power saving.

Lemma 6.20 Let \( \mathcal{Z} \subset \mathcal{Z}_1 \) be saturated with \( a_{11} \in \mathcal{Z}, a_{g+1,g+1} \notin \mathcal{Z} \) and \( N(\mathcal{L}(M), L, \mathcal{Z}, X) > 0 \). Suppose \( K_1 \subset \mathcal{Z} \). For any \( s \in T_\mathcal{Z}(L,Y) \), if
\[
I(K_1, s) \gg X^{-\delta}, \quad \text{and} \quad \#((s(YD) \times s(YD)) \cap W(\mathcal{Z})) \delta(s) \gg X^{n+1-\delta},
\]
then
\[
s_i \ll Y^{258g^3\delta} \quad \text{for} \quad i = g + 3, \ldots, n. \quad (52)
\]

To obtain a further saving, we need to use the \( |g| \)-invariant!

Lemma 6.21 Suppose \( M > X^\eta \) where \( \eta > 0 \) is some fixed constant. Let \( \mathcal{Z} \subset \mathcal{Z}_1 \) be saturated with \( a_{11} \in \mathcal{Z}, a_{g+1,g+1} \notin \mathcal{Z} \) and \( N(\mathcal{L}(M), L, \mathcal{Z}, X) > 0 \). Suppose \( K_1 \subset \mathcal{Z} \). Then for \( \delta < \min(\eta, 1)/(1355g^6) \) and any \( s \in T_\mathcal{Z}(L,Y) \) such that (51) and (52) hold, we have
\[
\#((s(YD) \times s(YD)) \cap \mathcal{L}(M)) \delta(s) \ll X^{n+1+514g^3\delta-1/2}.
\]
Therefore, by taking \( \delta = 64 \min(\eta, 1)/(1355n^6) \), we obtain the following result from Proposition 6.3 and Lemmas 6.15, 6.19, 6.20, and 6.21.

Proposition 6.22 Suppose \( M > X^\eta \) where \( \eta > 0 \) is some fixed constant. Let \( \mathcal{Z} \subset \mathcal{Z}_1 \) be saturated with \( a_{11} \in \mathcal{Z}, a_{g+1,g+1} \notin \mathcal{Z} \) and \( N(\mathcal{L}(M), L, \mathcal{Z}, X) > 0 \). Suppose \( K_1 \subset \mathcal{Z} \). Then
\[
N(\mathcal{L}(M), L, \mathcal{Z}, X) \ll X^{n+1-64 \min(\eta, 1)/(1355n^6)}.
\]

Theorem 6.11 then follows immediately from (39), Proposition 6.18, Proposition 6.22 and summing over the \( O(1) \) different possible \( \mathcal{Z} \)'s and the \( O(Y^e) \) different possible \( L \)'s.
6.3.3 Proofs of Lemmas 6.15, 6.16, 6.17, 6.19, 6.20 and 6.21.

We fix a saturated $Z \subset Z_1$ with $a_{11} \in Z$, $a_{g+1,g+1} \notin Z$ and $N(\mathcal{L}(M), L, Z, X) > 0$.

**Proof of Lemma 6.15:** Recall that $K_1 := \{a_{1,j} : 1 \leq j \leq g + 2\}$. Let $k \in \{1, \ldots, g\}$ be the largest integer such that $a_{kk} \notin Z$. Then, applying Lemma 6.14 to the saturated set $Z \cup K_1$, we have

$$
\left( \prod_{a \in Z \setminus K_1} \frac{w(\pi_k(a))}{w(a)} \right) \left( \prod_{1 \leq i < j \leq n \atop i+j>n+1} \frac{L_i}{L_j} \right) \delta(s) \ll \prod_{a \in K_1} \frac{w(\pi_k(a))}{w(a)} = \prod_{i=1}^{g} \delta_i^{-(n+1)(g+2)} \prod_{i=g+1}^{n-1} \delta_i^{-(n+1)(n-i)}.
$$

Hence, by Lemma 6.13, we obtain for any $s \in T_Z(L, Y)$,

$$
Y^{-\#Z} w(Z)^{-1} \left( \prod_{1 \leq i < j \leq n \atop i+j>n+1} \frac{L_i}{L_j} \right) \delta(s) \ll Y^{-\#Z} w(Z)^{-1} \left( \prod_{a \in Z \setminus K_1} w(\pi_k(a)) \right) \left( \prod_{1 \leq i < j \leq n \atop i+j>n+1} \frac{L_i}{L_j} \right) \delta(s)
$$

$$
\ll Y^{-\#(Z \cap K_1)} w(Z \cap K_1)^{-1} \left( \prod_{a \in Z \setminus K_1} \frac{w(\pi_k(a))}{w(a)} \right) \left( \prod_{1 \leq i < j \leq n \atop i+j>n+1} \frac{L_i}{L_j} \right) \delta(s)
$$

$$
= Y^{-\#(Z \cap K_1)} w(Z \cap K_1)^{-1} \prod_{i=1}^{g} \delta_i^{-(n+1)(g+2)} \prod_{i=g+1}^{n-1} \delta_i^{-(n+1)(n-i)}
$$

$$
= I(Z \cap K_1, s),
$$
as desired. □

Note that a direct computation yields

$$
I(K_1, s) = Y^{-(g+2)} \prod_{j=1}^{g+2} t_{n+1-j} t_j = Y^{-(g+2)} \frac{t_{g+1} t_{g+2}}{t_{n+1}},
$$

(53)

**Proof of Lemma 6.16:** Since $Z$ is saturated and $Z \cap K_1 \neq K_1$, we have $Z \cap K_1 = \{a_{11}, \ldots, a_{1j}\}$ for some $j = 1, \ldots, g + 1$. Since $a_{g+1,g+1}$ and $a_{1,g+2}$ do not belong to $Z$, we have for $s \in T_Z(L, Y)$:

$$
I(\{a_{11}, \ldots, a_{1,g+1}\}, s) = Y I(K_1, s) w(a_{1,g+2})
$$

$$
\ll Y^{g+1} I(K_1, s) w(a_{g+1,g+1}) w(a_{1,g+2})^g
$$

$$
\ll Y^{-1} \frac{1}{t_{g+1} t_{g+2} t_{n+1}}
$$

$$
\ll Y^{-1} s_n^{2},
$$
since the powers of the $s_i$'s in the third line are negative for $i < n$.

Similarly, for any $j = 1, \ldots, g$, we compute

$$
I(\{a_{11}, \ldots, a_{1j}\}, s) (Y w(a_{1,j+1}))^{j-1} \ll Y^{-1} s_n^2,
$$
as desired. □

**Proof of Lemma 6.17:** Suppose now $s_n \gg Y^{1/2-\delta}$. First note, that the inequality

$$
1 \ll Y^{g+1} w(a_{1,n}) w(a_{2,n-1}) \cdots w(a_{g+1,g+2}) = Y^{g+1} \prod_{i=1}^{n} s_i^{-i}
$$

(54)
implies that we have
\[ \prod_{j=1}^{n-1} s_j^{-1} \ll Y^{g+1} s_n^{-n} \ll Y^{n\delta}. \] (55)
Since each \( s_i \gg 1 \), we also have \( s_n \ll Y^{1/2} \) by (55). Hence
\[ t_1^{-1} \ll t_2^{-1} \ll \ldots \ll t_n^{-1} = s_n^{-n} \prod_{j=1}^{n-1} s_j^{-1} \ll Y^{-1/2+(n+1)\delta}; \quad t_{n+1}^{-1} = s_n^{-n} \prod_{j=1}^{n-1} s_j^{-1} \ll Y^{n/2+n\delta}. \]

Thus
\[ Yw(a_{ij}) = Yw(b_{ij}) = \frac{Y}{t_it_j} \ll \begin{cases} Y^{(2n+2)\delta} & \text{if } i \leq j \leq n, \\ Y^{(n+1)/2+(2n+1)\delta} & \text{if } i \leq n, j = n + 1, \\ Y^{n+1+2n\delta} & \text{if } i = j = n + 1. \end{cases} \]

Multiplying these weights together and applying Proposition 4.4 gives the estimate
\[ \#(s(Y\mathcal{D}) \times s(Y\mathcal{D}) \cap W_{n+1}(\mathbb{Z})) \ll Y^{(n+2)(n+1)+2n(n+2)\delta}. \] (56)

Meanwhile, in this region where \( s_n \geq X^{1/2-\delta} \), the quantity \( \delta(s) \) satisfies
\[ \delta(s) = \prod_{k=1}^{n} s_k^{-(n+1)k(n+1-k)} \ll s_n^{-n(n+1)} \ll Y^{-n(n+1)/2+n(n+1)\delta}. \] (57)

Multiplying the bounds in (56) and (57) together yields
\[ \#((s(Y\mathcal{D}) \times s(Y\mathcal{D})) \cap W_{n+1}(\mathbb{Z})) \delta(s) \ll Y^{(n+2)(n+1)+n(2n^2+9n+9)\delta}, \]
as desired. \( \square \)

**Proof of Lemma 6.19:** Suppose now \( K_1 \subset Z \subset K_1 \) and \( I(K_1, s) \gg X^{-\delta} \) for some \( s \in T_{Z}(L, Y) \).

We prove first that for any \( i = 1, \ldots, g - 1 \), we have
\[ Y^{-\delta} \ll \frac{s_i}{s_{n-i}} \ll Y^{\delta}. \] (58)

Indeed, since \( a_{j,n+1-j} \notin Z \) for all \( j \), we have from (53) that, for any \( k = 1, \ldots, g \),
\[ I(K_1, s) \ll I(K_1, s) Y^{g+2} w(a_{k,n+1-k}) w(a_{g+1,g+2})^2 \]
\[ \ll \frac{t_{g+1} t_{g+2}}{t_n t_{n+1}} \frac{t_g^2 t_{g+1}^2 t_{n-k+1}^2}{t_k^2 t_{n-k+1}^2} \]
\[ = \frac{t_1 \cdots t_g t_n \cdots t_{g+3}}{t_k^2 t_{n-k+1}^2} \]
\[ = \frac{t_1}{t_g} \cdots \frac{t_{g-1}}{t_g} \left( \frac{t_g}{t_k} \right)^g \frac{t_{g+4}}{t_{g+3}} \cdots \frac{t_n}{t_{g+3}} \left( \frac{t_{g+3}}{t_{n-k+1}} \right)^g. \]

Hence
\[ I(K_1, s) \ll \left( \prod_{i=1}^{g-1} \frac{s_i^{i(n+1)}}{(s_{n-i})^{i(n+1)}} \right)^{n+1} \left( \prod_{i=k}^{g-1} \frac{s_i^{i(n+1)}}{(s_{n-i})^{i(n+1)}} \right)^{-g(n+1)}. \]

38
Denote the product of the two factors in the final line by \( J_k \). Then
\[
\prod_{k=1}^{g} J_k = 1 \quad \text{and} \quad \frac{J_{i+1}}{J_i} = \left( \frac{s_i}{s_{n-i}} \right)^{g(n+1)} \quad \text{for } i = 1, \ldots, g-1.
\]

Since, by assumption, \( I(K_1, s) \gg X^{-\delta} \), we have \( J_k \gg Y^{-(n+1)\delta} \) for every \( k = 1, \ldots, g \). Therefore, for every \( i = 1, \ldots, g-1 \), we have
\[
\frac{s_i}{s_{n-i}} = \left( \frac{J_{i+1}}{J_i} \right)^{\frac{1}{g(n+1)}} = (J_1 \cdots J_{i-1} \cdot J_{i+1}^2 \cdot J_{i+2} \cdots J_g)^{\frac{1}{g(n+1)}} \gg Y^{-\delta};
\]
\[
\frac{s_i}{s_{n-i}} = \left( \frac{J_i}{J_{i+1}} \right)^{-\frac{1}{g(n+1)}} = (J_1 \cdots J_{i-1} \cdot J_i^2 \cdot J_{i+2} \cdots J_g)^{-\frac{1}{g(n+1)}} \ll Y^{\delta}.
\]

The claimed bound (58) follows.

By (54) and (58), we have
\[
s_g^g s_{g+1}^g s_{g+2}^g \ll Y^{g+1} \prod_{i=1}^{g-1} \left( \frac{s_i}{s_{n-i}} \right)^{-i} \prod_{i=g+3}^{n-1} s_i^{-n} \ll Y^{g+1 + \frac{3(g+1)\delta}{2}} \cdot R^{-n} \quad (59)
\]
where
\[
R = \prod_{i=g+3}^{n} s_i.
\]

We next prove the desired lower bounds:
\[
s_g \gg Y^{1/2 - (g/2)\delta} R^{-1}; \quad s_{g+1} \gg 1; \quad s_{g+2} \gg Y^{1/2 - \delta} R^{-1}. \quad (60)
\]

The bound \( s_{g+1} \gg 1 \) follows from the definition of \( T' \). For the bounds on \( s_g \) and \( s_{g+2} \), we use the assumption that \( a_{g+1,g+1} \notin \mathbb{Z} \) and the computation of \( I(K_1, s) \) in (53) to obtain
\[
I(K_1, s) \ll I(K_1, s) Y^{1/2} w(a_{g+1,g+1})^{1/2} = Y^{-(n+1)/2} t_{g+2} t_{n+1} = Y^{-(n+1)/2} s_{g+1} \frac{s_{g+1}^{n+1}}{g+2} R^{n+1},
\]
which along with \( I(K_1, s) \gg Y^{-(n+1)\delta} \) implies the desired lower bounds on \( s_{g+2} \); and
\[
I(K_1, s) \ll I(K_1, s) Y^{2} w(a_{g+1,g+1})^{2} = Y^{-g} \left( \frac{t_{g+2}}{g+1} \right)^{t_{n+1}} = Y^{-g} \left( \prod_{i=1}^{g} s_i^{3i} \right) s_{g+1}^{3(g+2)} \left( \prod_{i=g+2}^{n-1} s_i^{-2n-2+3i} \right)
\ll Y^{-g} \left( \prod_{i=1}^{g-1} s_i^{3i} \right)^{g-1} s_{g+1}^{3g} s_{g+1}^{g-3(g+2)} s_{g+2}^{-g} \left( \prod_{i=g+3}^{n-1} s_i^{-2n-2+3i} \right)^{2g}
\ll Y^{-g} \left( \prod_{i=1}^{g} s_i^{3i} \right)^{g(g+1)/2} s_{g+1}^{3g} s_{g+1}^{2g} R^{2g} \ll Y^{-g} \frac{3g}{2} \left( \prod_{i=1}^{g} s_i^{3i} \right) s_{g+1}^{3g} R^{3g},
\]
\[
\]
implying the desired lower bound on $s_g$, where in the last inequality we used the already-established lower bounds on $s_{g+1}$ and $s_{g+2}$.

The desired lower bounds for $s_g$, $s_{g+1}$, $s_{g+2}$ then follow by combining the upper bound on $s^g_{g+1}s^g_{g+2}$ in (59) and the individual lower bounds on $s_g$, $s_{g+1}$, $s_{g+2}$ in (60). The desired upper bound on $\mathcal{R}$ follows by comparing the upper bound on $s_g$ and the trivial lower bound $s_g \gg 1$. \qed

**Proof of Lemma 6.20:** Suppose $\mathcal{K}_1 \subset \mathcal{Z}$ and $s \in T_\mathcal{Z}(L,Y)$ satisfies (51). Then

$$
t_j^{-1} \leq \begin{cases} 
Y^{-(g+2)+5g^2\delta} \prod_{i=n-j+1}^{n} s_i^{n+1} & \text{for } j = 1, \ldots, g, \\
Y^{-1/2+20g^2\delta} & \text{for } j = g+1, g+2, \\
Y^{g+1+23g^2\delta} \prod_{i=j}^{n} s_i^{-(n+1)} & \text{for } j = g+3, \ldots, n+1,
\end{cases}
$$

(61)

where the upper bound on $\mathcal{R}$ also gives $t_j^{-1} \leq Y^{-1/2+20g^2\delta}$ for $j = 1, \ldots, g$. For $i, j \leq g+2$, we have $Y t_i^{-1} t_j^{-1} \leq Y^{40g^2\delta}$. For $j \geq g+3$ and $i \leq n-j+1$, we have $Y t_i^{-1} t_j^{-1} \leq Y^{28g^2\delta}$. Using (61) for the rest of the coordinates gives

$$
\#((s(YD) \times s(YD)) \cap W(\mathcal{Z})) \ll Y^{(n+1)(g+1)(g+4)+972g^4\delta} \left( \prod_{i=1}^{g} s_{g+2+i}^{-2i(n+3)(n+1)} \right).
$$

The Haar measure satisfies the following bound:

$$
\delta(s) = \prod_{k=1}^{n} s_k^{-(n+1)(k(n+1-k))} \ll Y^{-(g^2+3g+1)(n+1)+55g^4\delta} \left( \prod_{i=1}^{g} s_{g+2+i}^{2i(n+1)} \right).
$$

Hence

$$
\#((s(YD) \times s(YD)) \cap W(\mathcal{Z})) \delta(s) \ll Y^{(n+1)^2+1027g^4\delta} \left( \prod_{i=1}^{g} s_{g+2+i}^{-2i(n+1)} \right),
$$

(62)

Suppose now $\#((s(YD) \times s(YD)) \cap W(\mathcal{Z})) \delta(s) \gg X^{n+1-\delta}$. Then, for any $i = g+3, \ldots, n$,

$$
s_i \ll Y^{(1027g^4+2g+3)/(2(g-g-2)(g+3))} \ll Y^{1032g^4/(4g)} = Y^{258g^4\delta},
$$

as desired. \qed

**Proof of Lemma 6.21:** Suppose $M > X^\eta$ where $\eta > 0$ is some fixed constant. Suppose $\delta < \max(\eta,1)/1355g^6$. Suppose $\mathcal{K}_1 \subset \mathcal{Z}$ and $s \in T_\mathcal{Z}(L,Y)$ satisfies (51) and (52). We now impose the conditions $\det(B) = 0$ and $|q(A,B)| > M$ for any $(A,B) \in \mathcal{L}(M)$ to obtain a further saving for $\#((s(YD) \times s(YD)) \cap \mathcal{L}(M)) \delta(s)$.

The bound (52) on $s_g, \ldots, s_n$ gives $\mathcal{R} \ll Y^{258g^4\delta}$. Hence

$$
t_j^{-1} \leq \begin{cases} 
Y^{-(g+2)+1295g^2\delta}, & \text{for } j = 1, \ldots, g, \\
Y^{-1/2+20g^2\delta}, & \text{for } j = g+1, g+2, \\
Y^{g+1+23g^2\delta}, & \text{for } j = g+3, \ldots, n+1,
\end{cases}
$$

(63)

thus improving (61). In this case,

$$
Y t_{g}^{-1} t_{g+2}^{-1} \ll Y^{-(n+1)/2+1315g^2\delta}, \quad Y t_{g}^{-1} t_{n+1}^{-1} \ll Y^{1318g^2\delta},
$$

$$
Y t_{g+2}^{-1} t_{n+1}^{-1} \ll Y^{(n+1)/2+43g^2\delta}, \quad Y t_{n+1}^{-2} \ll Y^{n+1+46g^2\delta}.
$$
Since $\delta < 1/(1315g^4)$, we may assume that every $(A, B) \in (s(Y\mathcal{D}) \times s(Y\mathcal{D})) \cap W(\mathbb{Z})$ satisfies:

(a) The top left $g \times (g + 2)$-blocks of $A$ and $B$ are 0.

(b) The entries of the top right $g \times (g + 1)$ blocks of $A$ and $B$ are $O(Y^{1315g^5}\delta)$.

(c) The entries $a_{g+1,g+1}, a_{g+1,g+2}, a_{g+2,g+2}, b_{g+1,g+1}, b_{g+1,g+2},$ and $b_{g+2,g+2}$ are $O(Y^{40g^2}\delta)$.

(d) The entries $a_{g+1,j}, a_{g+2,j}, b_{g+1,j}$ and $b_{g+2,j}$ are $O(Y^{(n+1)/2 + 43g^2}\delta)$ for $g + 3 \leq j \leq n + 1$.

(e) The entries $a_{ij}$ and $b_{ij}$ are $O(Y^{n+1+46g^2}\delta)$ for $g + 3 \leq i, j \leq n + 1$.

Suppose now that $(A, B)$ is an element of $(s(Y\mathcal{D}) \times s(Y\mathcal{D})) \cap L(M)$. Then $f_{A,B} = xg(x, y)$, where $g(x, 1)$ is a degree $n$ polynomial with Galois group $S_n$.

**Lemma 6.23** Let $(A, B)$ be as above. If $b_{g+1,g+1} = b_{g+1,g+2} = b_{g+2,g+2} = 0$, then

$$|q|(A, B) \ll X^{1355g^6\delta}.$$

**Proof:** Let $v = x_0e_{g+1} + y_0e_{g+2}$ where $x_0, y_0 \in K$ are not both 0, where $K$ is some quadratic extension of $\mathbb{Q}$, and where $(x_0, y_0)$ is a solution to

$$a_{g+1,g+1}x^2 + a_{g+1,g+2}xy + a_{g+2,g+2}y^2 = 0. \quad (64)$$

We claim first that $v$ can be chosen in $\mathbb{Q}^{n+1}$. Indeed, $\text{Span}_K\{e_1, \ldots, e_g, v\}$ is a $(g + 1)$-dimensional subspace isotropic with respect to $A$ and $B$. Since $(A, B)$ is distinguished over $\mathbb{Q}$, the set of $(g + 1)$-dimensional common isotropic subspaces defined over any number field $L$ is in bijection with $J[2](L)$, where $J$ is the Jacobian of the hyperelliptic curve $y^2 = xg(x, 1)$ (which has a rational Weierstrass point at infinity), and $J[2](L)$ is in bijection with the factorizations of $xg(x, 1)$ over $L$. Since $g(x, 1)$ has Galois group $S_n$, it does not admit any factorization over any quadratic extension of $\mathbb{Q}$. Therefore, $J[2](K) = J[2](\mathbb{Q})$ with the only nontrivial element corresponding to the factorization $x \cdot g(x, 1)$. In other words, $\text{Span}_K\{e_1, \ldots, e_g, v\}$ is defined over $\mathbb{Q}$ and we may choose $v$ to be in $\mathbb{Q}^{n+1}$. In particular, the discriminant $a_{g+1,g+2}^2 - 4a_{g+1,g+1}a_{g+2,g+2} \in \mathbb{Z}$ is a square.

If $a_{g+1,g+1} \neq 0$, let

$$x_1 = -a_{g+1,g+2} + \sqrt{a_{g+1,g+2}^2 - 4a_{g+1,g+1}a_{g+2,g+2}}, \quad y_1 = 2a_{g+1,g+1}.$$

If $a_{g+1,g+1} = 0$, let $x_1 = 1, y_1 = 0$. Then $x_1, y_1$ are integers $\ll Y^{40g^2\delta}$, not both zero, and are solutions to (64). Let $x_0 = x_1 / \gcd(x_1, y_1)$ and $y_0 = y_1 / \gcd(x_1, y_1)$. Then there exist integers $x_2, y_2 \ll Y^{40g^2\delta}$ such that

$$\{e_1, \ldots, e_g, x_0e_{g+1} + y_0e_{g+2}; x_2e_{g+1} + y_2e_{g+2}; e_{g+3}, \ldots, e_n\}$$

forms an integral basis for $\mathbb{Z}^{n+1}$ such that the first $g + 1$ vectors generate a primitive lattice isotropic with respect to $A$ and $B$, and the first $g + 2$ vectors generate a primitive lattice isotropic with respect to $B$. That is, we compute the $|q|$-invariant of $(A, B)$ using this basis. When so expressed, the top right $(g + 1) \times (g + 2)$ blocks of the Gram matrices of $A$ and $B$ have the form

$$A^{\text{top}} = \begin{pmatrix} 0 & b & \cdots & b \\ \vdots & \ddots & \ddots & \vdots \\ 0 & b & \cdots & b \\ b & \cdots & \cdots & * \end{pmatrix}, \quad B^{\text{top}} = \begin{pmatrix} 0 & b & \cdots & b \\ \vdots & \ddots & \ddots & \vdots \\ 0 & b & \cdots & b \\ 0 & \cdots & \cdots & * \end{pmatrix},$$

41
Hence, to complete the proof of Lemma 6.21, it remains to prove that

\[ \text{consider the number of pairs } (\delta < \eta/(1355) \text{ by viewing} \]

for then we would have, using (62), that

\[ \text{it follows that } \delta \text{ is a nonzero polynomial, and entries labeled } "\sharp" \text{ are } O(Y^{12710g^6}) \text{, and entries labeled } "*" \text{ are } O(Y^{(n+1)/2 + 1438g^6}), \text{ where the top right coefficient } m' \text{ of } M_1 \text{ is the determinant of the top right } (g+1) \times (g+1) \text{ block } B' \text{ of } B_\text{top}, \text{ up to sign. Thus} \]

\[ |q|(A, B) = \frac{\|Q(A, B)\|}{\|\det(B')\|} = \frac{\|\det(M_1)\|}{|m'|} = |\det(M'_1)|, \]

where \( M'_1 \) is the bottom left \((g+1) \times (g+1)\) block of \( M_1 \). Since the coefficients of \( M'_1 \) are \( \ll Y^{2710g^6}, \) it follows that \( |q|(A, B) \ll X^{2710g^6(n+1)/n+1} \ll X^{1355g^6} \). \( \square \)

We now return to the proof of Lemma 6.21. For any \((A, B) \in (s(YD) \times s(YD)) \cap L(M), \) since \(|q|(A, B) > M > X^\eta, \) we may assume that \( b_{g+1,g+1}, b_{g+1,g+2}, \) and \( b_{g+2,g+2} \) are not all 0 since \( \delta < \eta/(1355g^6) \).

We now fix \( b_{ij} \) for \( 1 \leq i \leq g, g + 3 \leq j \leq n + 1, \) and \( i = g + 1, g + 2, j = g + 1, g + 2. \) We consider the number of pairs \((A, B) \in (s(YD) \times s(YD)) \cap L(M) \) with these prescribed coefficients by viewing \( \det(B) \) as a polynomial \( F \) in \( b_{ij} \) for \( g + 1 \leq i \leq n + 1 \) and \( g + 3 \leq j \leq n + 1. \) Note that all of these remaining coefficients have range at least

\[ Y^{(n+1)/2 + 43g^2} \prod_{i=g+1}^{n} s_i^{-(n+1)}. \]

Hence, to complete the proof of Lemma 6.21, it remains to prove that \( F \) is a nonzero polynomial, for then we would have, using (62), that

\[ \#((s(YD) \times s(YD)) \cap L(M)) \delta(s) \ll Y^{(n+1)^2 + 1027g^4} - (n+1)/2 - 43g^2 \prod_{i=1}^{g} s_i^{-(2i-1)(n+1)} \]

\[ \ll X^{n+1+514g^2/2}. \]

We may assume that the top right \( g \times (g+1) \) block of \( B \) has full rank, for otherwise the kernel of \( B \) would be isotropic with respect to \( A \) forcing \( \Delta(A, B) = 0 \) by Lemma 6.4. Hence we may also assume that the top right \( g \times (g+1) \) block of \( B \) equals \((I_g, 0), \) where \( I_g \) denotes the \( g \times g \) identity matrix. Then

\[ \det(B) = \det\begin{pmatrix} b_{g+1,g+1} & b_{g+1,g+2} & b_{g+1,n+1} \\ b_{g+2,g+1} & b_{g+2,g+2} & b_{g+2,n+2} \\ b_{n+1,g+1} & b_{n+1,g+2} & b_{n+1,n+1} \end{pmatrix}. \]

Since

\[ \begin{pmatrix} b_{g+1,g+1} & b_{g+1,g+2} \\ b_{g+2,g+1} & b_{g+2,g+2} \end{pmatrix} \neq 0, \]

we see that \( \det(B) \) is a nonzero polynomial in \( b_{g+1,n+1}, b_{g+2,n+1}, \) and \( b_{n+1,n+1}. \) \( \square \)
6.4 Bounding the number of distinguished elements in the deep cusp

In this subsection, we bound the number of elements with large \( q \)-invariant that lie in the deep cusp.

**Theorem 6.24** We have \( \mathcal{T}_X^{dcusp}(\mathcal{L}(M)) = O\left(\frac{X^{n+\frac{1}{2}+\frac{\kappa}{M}}}{} \log^2 X\right) \).

Recall from (39) that

\[
\mathcal{T}_X^{dcusp}(\mathcal{L}(M)) \ll \sum_L \sum_{Z:sg_{g+1} \in Z} N(\mathcal{L}(M), L, Z, X);
\]

here the first sum is over \( r \)-tuples \( L = (L_1, \ldots, L_r) \) with \( L_1 \leq L_2 \leq \cdots \leq L_n \) that partition the region \( \{(\mu_1, \ldots, \mu_r) \in [Y^{-\Theta_1}, Y^{\Theta_2}]^r : \mu_1 \leq \cdots \leq \mu_r\} \) into dyadic ranges, and the second sum is over saturated subsets \( Z \) of \( \mathcal{K} \cap \mathcal{M} \), where

\[
N(\mathcal{L}(M), L, Z, X) = \int_{T_2(Y, L)} \# \{(A, B) \in (s(YD) \times s(YD)) \cap \mathcal{L}(M) : B \in S(L, s)\} \delta(s) \, d^r s.
\]

The set \( S(L, s) \) is the union over \( \Lambda \in \Sigma(L, s) \) of \( S(\Lambda) \), where \( S(\Lambda) \) denotes the lattice of integral symmetric matrices whose row space is contained in \( \Lambda \otimes \mathbb{R} \), and \( \Sigma(L, s) \) denotes the set of primitive lattices \( \Lambda \in \mathbb{Z}^{n+1} \) of rank \( n \) such that the successive minima \( \mu_1, \ldots, \mu_n \) of \( s^{-1}(\Lambda) \) satisfy \( L_i \leq \mu_i < 2L_i \) for each \( i \in \{1, \ldots, n\} \). Finally recall from §6.1 and Proposition 5.4 that

\[
w(\ell_{g+1} t_{g+1}^2) = L_{g+1}^2 \leq c_1 \ell t_{g+1}^2 = c_1 Y w(a_{g+1} t_{g+1}) < c_1 c_{g+1} t_{g+1}^2 < c_{g+1}^2
\]

for every \( s \in T_2(Y, L) \). Hence we may assume that \( c_{g+1} t_{g+1} \in \mathbb{Z} \).

The deep cusp contains \( \times X^{n+1} \) elements, and we obtain a saving because the elements we are counting have \( q \)-invariant greater than \( M \). To make use of this condition, we require an upper bound on the size of the \( |q| \)-invariant of elements in \( (s(YD) \times s(YD)) \cap \mathcal{L}(1) \). To accomplish this, we have the following preliminary result.

**Lemma 6.25** Let \( (A, B) \in (YD \times YD) \cap W_0(\mathbb{R}) \) be such that \( \Delta(A, B) > X^{2n-2-\kappa} \). Denote the top right \( (g + 1) \times (g + 2) \) block of \( B \) by \( B'^{top} \). Then

\[
\det(B'^{top}(B'^{top})^t) \gg Y^{2(g+1)-(n+1)\kappa}.
\]

**Proof:** Let \( (A', B') = Y^{-1}(A, B) \in (\mathcal{D} \times \mathcal{D}) \cap W_0(\mathbb{R}) \). Then it suffices to prove that

\[
\det(B'^{top}(B'^{top})^t) \gg Y^{-(n+1)\kappa}.
\]

Since \( |\Delta(A, B)| > X^{2n-2-\kappa} \), we have \( |\Delta(A', B')| > X^{-\kappa} \). By Proposition 3.6, there is a polynomial \( P \in \mathbb{Z}[W_0] \) such that

\[
\Delta(A'', B'') = P(A'', B'') \det(B''^{top}(B''^{top})^t)
\]

for any \( (A'', B'') \in W_0(\mathbb{R}) \). Since \( (A', B') \in \mathcal{D} \times \mathcal{D} \), which is an absolutely bounded region, we have \( |P(A', B')| \ll 1 \). Hence

\[
\det(B'^{top}(B'^{top})^t) = \frac{\Delta(A', B')}{P(A', B')} \gg X^{-\kappa}
\]

as desired. \( \Box \)

Next, we have the following upper bound on the \(|q|\)-invariant.
\textbf{Proposition 6.26} Let $Z \subseteq Z_1$ be a saturated set containing $a_{g+1,g+1}$ and $\ell_{g+1,g+1}$. Let $L = (L_1, \ldots, L_n)$ be a sequence of nondecreasing positive real numbers. Then for any $s \in T_Z(L, Y)$ and $(A, B) \in (s(YD) \times s(YD)) \cap L(1)$, we have

$$|q|(A, B) \ll Y^{(g+1)^2 + \frac{n+1}{4}} \prod_{i=1}^{g+1} L_i, \quad (66)$$

\textbf{Proof:} Suppose $(A, B) \in (s(YD) \times s(YD)) \cap L(1)$. Since $a_{g+1,g+1} \in Z$, we have $(A, B) \in W_0(Z)$. By Lemma 6.4, $\ker(B)$ is 1-dimensional and does not lie inside $\text{Span}\{e_1, \ldots, e_{g+1}\}$ as this $(g + 1)$-plane is isotropic with respect to $A$. Let $w_1 \in \text{Span}_Z\{e_{g+2}, \ldots, e_{n+1}\}$ be a primitive vector so that $\{e_1, \ldots, e_{g+1}, w_1\}$ forms a basis for the primitive lattice in $\text{Span}_R\{e_1, \ldots, e_{g+1}\} + \ker(B)$. Complete $w_1$ to an integral basis $\{w_1, \ldots, w_{g+2}\}$ for $\text{Span}_Z\{e_{g+2}, \ldots, e_{n+1}\}$. We can now use the integral basis

$$\{e_1, \ldots, e_{g+1}, w_1, \ldots, w_{g+2}\} \quad (67)$$

of $Z^{n+1}$ to compute the $|q|$-invariant of $(A, B)$, as the first $g + 1$ vectors generate a primitive lattice isotropic with respect to $A$ and $B$, and the first $g + 2$ vectors generate a primitive lattice isotropic with respect to $B$. Note also that with respect to the standard inner product on $\mathbb{R}^{n+1}$, since $w_1 \in \text{Span}_R\{e_1, \ldots, e_{g+1}\} + \ker(B)$, we have

$$w_1 \perp (\text{Span}_R\{e_{g+2}, \ldots, e_{n+1}\} \cap C(B)) \quad (68)$$

where $C(B)$ denotes the column space of $B$.

Let $A'$ and $B'$ be the Gram matrices of the quadratic forms defined by $A$ and $B$ with respect to this new basis (67). Since the first $g + 1$ vectors of this basis are part of the standard basis, we see that $(A, B)$ and $(A', B')$ are $G_0(Z)$-equivalent, where $G_0$ is defined in §3.1. Hence

$$|Q|(A', B') = |Q|(A, B) \ll Y^{(g+1)(g+2)} \prod_{k=1}^{g+1} t_k^{-1}. \quad (69)$$

Let $B''$ denote the top right $(g + 1) \times (g + 1)$ block of $B'$. Then, by the definition of $q$, we have

$$|q|(A, B) = |q|(A', B') = \frac{|Q|(A', B')}{|\det(B'')|} \ll \frac{1}{|\det(B'')|} Y^{(g+1)(g+2)} \prod_{k=1}^{g+1} t_k^{-1}. \quad (69)$$

We now work towards proving a lower bound on $|\det(B'')|$. Let $p_1$ (resp., $p_2$) denote the projection of $\mathbb{R}^{n+1}$ onto the first $g + 1$ coefficients (resp., the last $g + 2$ coefficients). Let $B'^{\top}$ (resp., $B^{\top}$) denote the top right $(g + 1) \times (g + 2)$ block of $B$ (resp., $B'$). Then by (68), we have $B'^{\top}p_2(w_1) = 0$. Consider the following two $(g + 2) \times (g + 2)$ matrices in block form:

$$B^* = \begin{pmatrix} B'^{\top} \\ p_2(w_1)^t \end{pmatrix}, \quad \gamma = \begin{pmatrix} p_2(w_1) & \cdots & p_2(w_{g+2}) \end{pmatrix}.$$

Then

$$B^* \gamma = \begin{pmatrix} B'^{\top} \\ |w_1|^2 \end{pmatrix}.$$

Let $\Lambda_2$ denote the rank $g + 1$ lattice in $\mathbb{Z}^{g+2}$ spanned by the rows of $B'^{\top}$. Then $|\det(B^*)| = d(\Lambda_2) |w_1|$. Since $\{p_2(w_1), \ldots, p_2(w_{g+2})\}$ is an integral basis for $\mathbb{Z}^{g+2}$, we have $\det \gamma = \pm 1$ and so

$$|\det(B'')| = \frac{|\det(B^*)\det \gamma|}{|w_1|^2} = \frac{d(\Lambda_2) |w_1| \cdot 1}{|w_1|^2} = \frac{d(\Lambda_2)}{|w_1|}.$$
We now use the fact that \( B \in \mathcal{S}(L, s) \). This means that the row span of \( B \) lies in an \( n \)-dimensional primitive lattice \( \Lambda \subset \mathbb{Z}^{n+1} \) with basis of the form \( \{s\ell_1, \ldots, s\ell_n\} \) where \( L_i \leq |\ell_i| < 2L_i \) and \( \{\ell_1, \ldots, \ell_n\} \) are reduced. By assumption, \( \ell_{g+1, g+1} \in \mathbb{Z} \), and hence \( \ell_{i,j} \in \mathbb{Z} \) for all \( i \leq g + 1 \) and \( j \leq g + 1 \). Thus the first \( g + 1 \) coefficients of \( s\ell_1, \ldots, s\ell_{g+1} \) are all 0, and \( \{s\ell_1, \ldots, s\ell_{g+1}\} \) forms an integral basis of a primitive lattice \( \Lambda_1 \) of rank \( g + 1 \) in \( \text{Span}_{\mathbb{R}} \{e_{g+2}, \ldots, e_{n+1}\} \). By (68), \( w_1 \) is a primitive vector in \( \text{Span}_{\mathbb{R}} \{e_{g+2}, \ldots, e_{n+1}\} \) orthogonal to \( \Lambda_1 \). Hence

\[
|w_1| = d(\Lambda_1).
\]

By (68), we have \( \text{Span}_{\mathbb{R}} \{e_{g+2}, \ldots, e_{n+1}\} \cap C(B) \neq \text{Span}_{\mathbb{R}} \{e_{g+2}, \ldots, e_{n+1}\} \), and so

\[
\text{Span}_{\mathbb{R}} \{e_{g+2}, \ldots, e_{n+1}\} \cap C(B) = \Lambda_1 \otimes \mathbb{R}.
\]

In particular, since \( \Lambda_1 \) is primitive, the first \( g + 1 \) columns of \( B \) belong to \( \Lambda_1 \). That is, there is a \( (g + 1) \times (g + 1) \) matrix \( C \) (with integer coefficients) such that

\[
B^{\text{top}} = C \begin{pmatrix} p_2(s\ell_1)^t \\ \vdots \\ p_2(s\ell_{g+1}) \end{pmatrix}
\]

and so

\[
|\det(C)| = \frac{d(\Lambda_2)}{d(p_2(\Lambda_1))} = \frac{d(\Lambda_2)}{|w_1|} = |\det(B'')|.
\]  (70)

To obtain a lower bound on \( |\det(C)| \), we write

\[
s = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}
\]  with

\[
A_1 = \begin{pmatrix} t_1^{-1} & & \\ & \ddots & \\
 & & t_{g+1}^{-1} \end{pmatrix}, \quad A_2 = \begin{pmatrix} t_1^{-1} \\ \vdots \\ t_{g+1}^{-1} \end{pmatrix}.
\]

Let \( M^{\text{top}} \) denote the \( (g + 1) \times (g + 2) \) matrix with rows \( p_2(\ell_1)^t, \ldots, p_2(\ell_{g+1})^t \). Then

\[
CM^{\text{top}}A_2 = B^{\text{top}}.
\]

Consider the pair \( (A_0, B_0) := s^{-1}(A, B) \in (YD \times YD) \cap W_{0,n+1}(\mathbb{R}) \) satisfying

\[
|\Delta(A_0, B_0)| = |\Delta(A, B)| > X^{2n-2-\kappa}
\]

since \( (A, B) \in \mathcal{L}(1) \). The top right \( (g + 1) \times (g + 2) \) block \( B_0^{\text{top}} \) of \( B_0 \) satisfies

\[
A_1B_0^{\text{top}}A_2 = B^{\text{top}},
\]

and so

\[
CM^{\text{top}} = A_1B_0^{\text{top}}.
\]

The rows of \( M^{\text{top}} \) form a reduced basis for a lattice \( \Lambda_3 \subset \mathbb{Z}^{g+2} \) with \( L_i \leq |p_2(\ell_i)| < 2L_i \). Thus

\[
\det(B'')^2 = \det(C)^2 = \frac{\det(A_1B_0'')(B_0^{\text{top}})^tA_1}{\det(M^{\text{top}}(M^{\text{top}})^t)} \gg \frac{t_1^{-2} \cdots t_{g+1}^{-2}}{L_1^2 \cdots L_{g+1}^2} \det(B_0'')(B_0^{\text{top}})^t.
\]  (71)
By Equations (69) and (69),
\[ |q|(A, B) \ll Y^{(g+1)(g+2)} \frac{L_1 \cdots L_g}{\sqrt{\det(B_0^{\text{top}}(B_0^{\text{top}})^t)}}. \]
The result now follows from Lemma 6.25. □

**Proof of Theorem 6.24:** We write
\[ I_X^{\text{dexp}}(L(M)) \ll \sum_L \sum_Z N(L(M), L, Z, X), \tag{72} \]
and obtain upper bounds on \( N(L(M), L, Z, X) \) for each \( Z \subset Z_1 \) with \( a_{g+1, g+1} \ell_{g+1, g+1} \in \mathbb{Z} \). Fix such a set \( Z \) with \( N(L(M), L, Z, X) > 0 \) and an element \( s \in T_Z(L, Y) \). Then
\[ (s(YD) \times s(YD)) \cap L(M) \subset (s(YD) \cap S(Z)) \times (s(YD) \cap S(L, s)). \]

We begin by bounding the number of elements in \( \#(s(YD) \cap S(Z)) \). Let \( K_{\text{dist}} := \{ a_{ij} \mid 1 \leq i \leq j \leq g + 1 \} \). By assumption, \( K_{\text{dist}} \) is a subset of \( Z \cap K \). Define \( \pi_K : Z_1 \cap K \to \mathbb{K} \setminus Z_1 \) by
\[ \pi_K(a_{ij}) := a_{n+1-j, j}. \]
This agrees with the \( \pi_k \) as defined in §6.3.1 when restricted to \( K \). For any \( \alpha \in Z_1 \cap K \), we have \( Y w(\pi_K(\alpha)) \gg 1 \) and \( w(\pi_K(\alpha)) \gg w(\alpha) \). For any \( a_{ij} \in (Z_1 \cap K) \setminus K_{\text{dist}} \), we have \( i < g + 1 < j \). Thus
\[ \prod_{\alpha \in (Z_1 \cap K) \setminus K_{\text{dist}}} \frac{w(\pi_K(\alpha))}{w(\alpha)} \ll \prod_{\alpha \in (Z_1 \cap K) \setminus K_{\text{dist}}} \frac{w(\pi_K(\alpha))}{w(\alpha)} = \prod_{1 \leq i < g + 1 < j \leq n} \frac{t_i}{t_{n+1-j}} \prod_{1 \leq i \leq j \leq g + 1} \frac{t_i}{t_j}. \]
Therefore,
\[ \#(s(YD) \cap S(Z)) \ll Y^{(n+1)(n+2)/2 - \#(Z \cap K)} \prod_{\alpha \in Z \cap K} \frac{1}{w(\alpha)} \]
\[ \ll Y^{(n+1)(n+2)/2 - \#(Z \cap K)} \left( \prod_{\alpha \in K_{\text{dist}}} \frac{1}{w(\alpha)} \right) \left( \prod_{\alpha \in (Z_1 \cap K) \setminus K_{\text{dist}}} \frac{Y w(\pi_q(\alpha))}{w(\alpha)} \right) \]
\[ \ll \frac{Y^{(n+1)(n+2)/2}}{Y \#K_{\text{dist}}} \left( \prod_{1 \leq i < j \leq g + 1} t_i t_j \right) \left( \prod_{1 \leq i \leq j \leq g + 1} \frac{t_i}{t_j} \right) \]
\[ = \frac{Y^{(n+1)(n+2)/2}}{Y^{(g+1)(g+2)/2}} (t_1 \cdots t_{g+1})^{g+2} \left( \prod_{1 \leq i \leq j \leq g + 1} \frac{t_i}{t_j} \right). \tag{73} \]

We now obtain an upper bound on \( \#(s(YD) \cap S(L, s)) \). Recall that
\[ \#(s(YD) \cap S(L, s)) = \sum_{\Lambda \in \Sigma(L, s)} \#(YD \cap s^{-1} S(\Lambda)). \tag{74} \]
Let \( \Lambda \in \Sigma(L, s) \) be a lattice such that \( s^{-1}(\Lambda) \) has reduced basis \( \{ \ell_1, \ldots, \ell_n \} \) with \( L_i \leq |\ell_i| < 2L_i \) for each \( i = 1, \ldots, n \). Suppose there exists \( (A, B) \in (s(YD) \times s(YD)) \cap L(M) \) with \( B \in s(YD) \cap S(\Lambda) \). By Proposition 6.26,
\[ M \ll Y^{(g+1)^2 + n+1} \prod_{i=1}^{g+1} L_i. \tag{75} \]

46
Recall also from the proof of Proposition 6.26 that
\[ \text{Span}_\mathbb{R}\{e_{g+2}, \ldots, e_{n+1}\} \cap C(B) = \text{Span}_\mathbb{R}\{s\ell_1, \ldots, s\ell_{g+1}\}. \]

Hence
\[ \text{Span}_\mathbb{R}\{e_{g+2}, \ldots, e_{n+1}\} \cap \text{Span}_\mathbb{R}\{s\ell_{g+2}, \ldots, s\ell_n\} = \{0\}. \]

It follows that the set \( \{p_1(s\ell_{g+2}), \ldots, p_1(s\ell_n)\} \), and thus the set \( \{p_1(\ell_{g+2}), \ldots, p_1(\ell_n)\} \), are both linearly independent. There then exist vectors \( v_{g+2}, \ldots, v_n \in \text{Span}_\mathbb{R}\{e_1, \ldots, e_{g+1}\} \) such that
\[ (v_{g+2} \quad \cdots \quad v_n)^t (\ell_{g+2} \quad \cdots \quad \ell_n) = I_{g+1} \]
is the identity matrix. Let \( B' \in s(Y\mathcal{D}) \cap S(\Lambda) \) be any element and write
\[ s^{-1}B' = \sum_{1 \leq i \leq j \leq n} \beta_{ij} \ell_i \ast \ell_j, \]
where \( \ell_i \ast \ell_j \) is as defined in (22). Then for \( g + 2 \leq i \leq j \leq n \), since \( v_i, v_j \perp \ell_1, \ldots, \ell_{g+2}, \) we have
\[ v_i^t(s^{-1}B')v_j = \begin{cases} 2\beta_{ij} & \text{if } i \neq j, \\ \beta_{ii} & \text{if } i = j. \end{cases} \]
Since the top left \((g + 1) \times (g + 1)\) block of \( B' \in s(Y\mathcal{D}) \cap S(\Lambda) \) is 0, the same is true for \( s^{-1}B' \).

Hence \( \beta_{ij} = 0 \) whenever \( g + 2 \leq i \leq j \leq n \). In other words,
\[ Y\mathcal{D} \cap s^{-1}S(\Lambda) \subset \text{Span}_\mathbb{R}\{\ell_i \ast \ell_j \mid 1 \leq i \leq j \leq n \text{ and } i \leq g + 1\}. \]

By Proposition 5.2, we have
\[ \#(Y\mathcal{D} \cap s^{-1}S(\Lambda)) \ll \prod_{1 \leq i \leq j \leq n} \frac{Y}{L_i L_j} \]
\[ \ll \left( \prod_{1 \leq i \leq j \leq n} \frac{Y}{L_i L_j} \right) \prod_{1 \leq i < j \leq n \atop i \leq g + 1} \left( \frac{L_i L_j}{Y^{(g+1)(g+2)/2}} \prod_{1 \leq i \leq j \leq n \atop i \leq g + 1} \frac{L_i}{L_{n+1-j}} \right) \]
\[ \ll \frac{Y^{n(n+1)/2}}{(L_1 \cdots L_n)^{n+1}} \frac{(L_{g+2} \cdots L_n)^{g+2}}{Y^{(g+1)(g+2)/2}} \prod_{1 \leq i \leq j \leq n \atop i \leq g + 1} \frac{L_i}{L_{n+1-j}} \]
\[ \ll \frac{Y^{n(n+1)/2}}{(L_1 \cdots L_n)^{n+1}} \frac{(L_{g+2} \cdots L_n)^{g+2}}{Y^{(g+1)(g+2)/2}} \prod_{1 \leq i < j \leq n \atop i \leq g + 1} \frac{L_i}{L_j}, \quad (76) \]

where the second bound follows since \( L_{n+1-j}L_j \ll Y \) for all \( j \) by Proposition 5.4; and the last bound follows because the map from \( \{(i, j) : 1 \leq i \leq j \leq n \text{ and } i \leq g + 1\} \) to \( \{(k, \ell)\} \) sending \((i, j)\) to \((n+1-j, i)\) is one-to-one with its image contained within the set of pairs \((k, \ell)\) with \( k < \ell \leq g+1\), and because the \( L_i \)'s are nondecreasing.

To obtain a bound on the size of \( \Sigma(L, s) \), we use (42):
\[ \#\Sigma(L, s) \ll (L_1 L_2 \cdots L_n)^{n+1} \left( \prod_{1 \leq i < j \leq n} \frac{L_i}{L_j} \right) \left( \prod_{\alpha \in \mathbb{Z} \cap \mathcal{M}} \frac{1}{\omega(\alpha)} \right). \]

47
Let $\mathcal{M}_{\text{dist}} = \{\ell_{i,j} \mid 1 \leq i \leq g+1, 1 \leq j \leq g+1\}$. Recall that elements $\ell_{i,j} \in \mathbb{Z}_1$ satisfy $i+j \leq n+1$. Hence for any $\ell_{i,j} \in (\mathbb{Z}_1 \cap \mathcal{M}) \setminus \mathcal{M}_{\text{dist}}$, exactly one of $i$ and $j$ is $\leq g+1$. Define

$$
\pi_{\mathcal{M}} : (\mathbb{Z}_1 \cap \mathcal{M}) \setminus \mathcal{M}_{\text{dist}} \to \mathcal{M}
$$

$$
\pi_{\mathcal{M}}(\ell_{i,j}) = \begin{cases} 
\ell_{i,n+2-i} & \text{if } i \leq g+1; \\
\ell_{n+1-j,j} & \text{if } j \leq g+1.
\end{cases}
$$

We claim that the image of $\pi_{\mathcal{M}}$ is disjoint from $\mathbb{Z}$. Indeed, when $i \leq g+1$, we have $\pi_{\mathcal{M}}(\ell_{i,j}) \notin \mathbb{Z}_1$, and when $j \leq g+1$, we have $\pi_{\mathcal{M}}(\ell_{i,j}) \notin \mathbb{Z}$ by Lemma 6.5 and the fact that $a_{g+1,g+1} \in \mathbb{Z}$. Thus $w(\pi_{\mathcal{M}}(\alpha)) \gg 1$ and $w(\pi_{\mathcal{M}}(\alpha)) \gg w(\alpha)$ for every $\alpha \in (\mathbb{Z}_1 \cap \mathcal{M}) \setminus \mathcal{M}_{\text{dist}}$. It follows that

$$
\prod_{\alpha \in \mathcal{M}} \frac{1}{w(\alpha)} \ll \left( \prod_{\ell \in \mathcal{M}_{\text{dist}}} \frac{1}{w(\ell)} \right) \left( \prod_{\ell \in \mathcal{M}_{\text{dist}}} \frac{w(\pi(\ell))}{w(\ell)} \right)
$$

$$
\ll \frac{(t_1 \cdots t_{g+1})^{g+1}}{(L_1 \cdots L_{g+1})^{g+1}} \left( \prod_{g+2 \leq i < j \leq n+1} \frac{t_i}{t_j} \right) \left( \prod_{g+2 \leq i \leq j \leq n+1} \frac{L_i}{L_j} \right),
$$

so that

$$
\# \Sigma(L,s) \ll \frac{(t_1 \cdots t_{g+1})^{g+1}}{(L_1 \cdots L_{g+1})^{g+1}} \left( \prod_{g+2 \leq i < j \leq n+1} \frac{t_i}{t_j} \right) \left( \prod_{g+2 \leq i \leq j \leq n+1} \frac{L_i}{L_j} \right).
$$

Combining (74), (76), and (77) and the identity

$$
\left( \prod_{1 \leq i < j \leq n+1} \frac{L_j}{L_i} \right) \left( \prod_{i=1}^{n} \prod_{j=i}^{n} \frac{L_i}{L_j} \right) = \frac{(L_1 \cdots L_{g+1})^{g+1}}{(L_{g+2} \cdots L_n)^{g+1}}
$$

now yields

$$
\#(s(YD) \cup S(L,s)) \ll \frac{Y^{n(n+1)/2}}{(L_1 \cdots L_n)^{n+1}} \frac{(L_{g+2} \cdots L_n)^{g+2}}{Y^{(g+1)(g+2)/2}} \left( \prod_{1 \leq i < j \leq g+1} \frac{L_j}{L_i} \right)
$$

$$
\cdot \left( \frac{(t_1 \cdots t_{g+1})^{g+1}}{(L_1 \cdots L_{g+1})^{g+1}} \left( \prod_{g+2 \leq i < j \leq n+1} \frac{t_i}{t_j} \right) \left( \prod_{g+2 \leq i \leq j \leq n+1} \frac{L_i}{L_j} \right) \right)
$$

$$
= \frac{Y^{n(n+1)/2}}{Y^{(g+1)(g+2)/2}} \left( \frac{(t_1 \cdots t_{g+1})^{g+1}}{(L_1 \cdots L_{g+1})^{g+1}} \left( \prod_{g+2 \leq i < j \leq n+1} \frac{t_i}{t_j} \right) \left( \prod_{g+2 \leq i \leq j \leq n+1} \frac{Y}{L_j L_{n+1-j}} \right) \right)
$$

$$
= \frac{Y^{n(n+1)/2+g+1}}{Y^{(g+1)(g+2)/2}} \left( \frac{(t_1 \cdots t_{g+1})^{g+1}}{(L_1 \cdots L_{g+1})^{g+1}} \prod_{g+2 \leq i < j \leq n+1} \frac{t_i}{t_j} \right),
$$

where the fourth line follows since $L_{n+1-j}L_j \ll Y$ for all $j$ by Proposition 5.4. Finally, note that

$$
(t_1 \cdots t_{g+1})^{2g+3} \prod_{1 \leq i < j \leq g+1} \frac{t_i}{t_j} \prod_{g+2 \leq i < j \leq n+1+1} \frac{t_i}{t_j} = \frac{(t_1 \cdots t_{g+1})^{g+2}}{(t_{g+2} \cdots t_{n+1})^{g+1}} \prod_{1 \leq i < j \leq g+1} \frac{t_i}{t_j} \prod_{g+2 \leq i < j \leq n+1} \frac{t_i}{t_j}
$$

$$
= \prod_{1 \leq i < j \leq n+1} \frac{t_i}{t_j} = \delta(s)^{-1}.
$$

48
Therefore, combining (73), (75) and (78) gives
\[
N(\mathcal{L}(M), L, Z, X) \ll \int_{T_x(L^*Y)} \#(s(\mathcal{D}) \cap S(Z)) \cdot \#(s(\mathcal{D}) \cap S(L, s)) \delta(s)d^x s \\
\ll \frac{Y^{(n+1)^2-(g+1)^2}}{L_1 \cdots L_{g+1}} \int_{T_x(L^*Y)} d^x s \\
\ll \frac{X^{(n+1)+\frac{\delta}{\kappa}}}{M} \log^2 Y.
\]
Theorem 6.24 now follows immediately from (72) by summing over the \(O(1)\) different possible \(Z\)’s and the \(O(\log^2 Y)\) different possible \(L\)’s. \(\Box\)

6.5 Proof of the main uniformity estimates

Proof: Case (a) of Theorem 5 follows from an application of the quantitative version of the Ekedahl geometric sieve developed in [3, Theorem 3.3]. Case (b) follows from (14), (16), and Theorem 4.1.

We now use the results of this section to prove the most intricate case, namely, Case (c). For \(m \leq X^{1/2}\), we have the immediate bound
\[
\#\{f \in \mathcal{W}_m : H(f) < X\} \ll X^{n+1}/m^2.
\]
Using this bound for \(m \leq X^{1/2}\), we may assume that \(M > X^{1/2}\). We first note that from (34), (36), and Lemmas 6.1 and 6.2, we have
\[
\# \bigcup_{m > M \atop m \text{ squarefree}} \{f \in \mathcal{W}_m^{(2)} : H(f) < X\} \ll \mathcal{I}_X(\mathcal{L}(\sqrt{M})) + \frac{X^{n+1}}{\sqrt{M}} + X^n + X^{n+1-\frac{\delta}{2n-2}}.
\]
Applying Theorems 6.6, 6.11, and 6.24, we obtain
\[
\mathcal{I}_X(\mathcal{L}(\sqrt{M})) = \mathcal{I}_X^{\text{main}}(\mathcal{L}(\sqrt{M})) + \mathcal{I}_X^{\text{scusp}}(\mathcal{L}(\sqrt{M})) + \mathcal{I}_X^{\text{dcusp}}(\mathcal{L}(\sqrt{M})) \\
\ll X^{n+1-1/(10n)} + X^{n+1-1/(88n^6)} + \frac{X^{n+1+\frac{\delta}{\kappa}}}{\sqrt{M}} \log^2 n X.
\]

Setting \(\kappa = (2n-2)/(88n^6)\) yields the desired result. \(\Box\)

Theorem 5 has the following immediate consequence. For a positive squarefree integer \(m\), let \(\mathcal{W}_m := \mathcal{W}_m^{(1)} \cup \mathcal{W}_m^{(2)}\).

Corollary 6.27 For a positive integer \(N \geq 3\), and positive real numbers \(M, X\), we have
\[
\sum_{m > M \atop m \text{ squarefree}} \#\{f \in \mathcal{W}_m : H(f) < X\} \ll \epsilon \frac{X^{n+1+\delta_n+\epsilon}}{M^{\delta_n}} + X^{n+1-\eta_n+\epsilon},
\]
where \(\delta_n = 1/2, \xi_n = 0, \eta_n = 1/(5n)\) when \(n\) is odd and \(\delta_n = 1/3, \xi_n = 1/(88n^5), \eta_n = 1/(88n^6)\) when \(n\) is even.

Proof: Suppose \(f \in \mathcal{W}_m\) for some \(m > M\). When \(n\) is odd, this implies that \(f\) belongs to either \(\mathcal{W}_m^{(1)}\) or \(\mathcal{W}_m^{(2)}\) for some \(m_1 > \sqrt{M}\). When \(n\) is even, \(f\) belongs either to \(\mathcal{W}_m^{(1)}\) for \(m_2 > M^{1/3}\) or to \(\mathcal{W}_m^{(2)}\) for \(m_3 > M^{2/3}\). A direct application of Theorem 5 now yields the result. \(\Box\)
7 Proofs of the main results

We begin by proving a more general form of Theorem 6. Let \( N \) be a positive squarefree integer, and for each \( p \mid N \), let \( \Sigma_p \subset V_n(\mathbb{Z}/p^2\mathbb{Z}) \) be a nonempty subset. Denote the collection \( (\Sigma_p)_{p\mid N} \) by \( \Sigma \). Let \( V_n(\Sigma) \) be the set of all \( f \in V_n(\mathbb{Z}) \) such that the reduction of \( f \) modulo \( p^2 \) lies in \( \Sigma_p \) for all \( p \mid N \). For \( p \mid N \), let \( \alpha_n(\Sigma, p) \) (resp., \( \beta_n(\Sigma, p) \)) denote the density of elements \( f \in V_n(\mathbb{Z}) \) such that \( p^2 \not\mid \Delta(f) \) (resp., \( R_f \) is maximal at \( p \)) and such that the reduction of \( f \) modulo \( p^2 \) lies in \( \Sigma_p \).

For \( p \nmid N \), simply set \( \alpha_n(\Sigma, p) = \alpha_n(p) \) and \( \beta_n(\Sigma, p) = \beta_n(p) \). Finally, define

\[
\alpha_n(\Sigma) = \prod_p \alpha_n(\Sigma, p); \quad \beta_n(\Sigma) = \prod_p \beta_n(\Sigma, p).
\]

We are now ready to carry out our sieve.

**Theorem 7.1** We have,

\[
\begin{align*}
\# \{ f \in V_n(\Sigma) : H(f) < X \text{ and } \Delta(f) \text{ squarefree} \} &= \alpha_n(\Sigma)(2X)^{n+1} + \mathcal{O}(\mathcal{E}(X, N, \epsilon)), \\
\# \{ f \in V_n(\Sigma) : H(f) < X \text{ and } R_f \text{ maximal} \} &= \beta_n(\Sigma)(2X)^{n+1} + \mathcal{O}(\mathcal{E}(X, N, \epsilon))
\end{align*}
\]

where the error term is given by

\[
\mathcal{E}(X, N, \epsilon) := X^{n+1-\eta_n+\epsilon} + N^2X^{n+3(\eta_n+\xi_n)+\epsilon} + N^{2n+2}X^{6(n+3)(\eta_n+\xi_n)+\epsilon},
\]

where \( \eta_n = 1/(5n) \), \( \xi_n = 0 \) when \( n \) is odd, and \( \eta_n = 1/(88n^6) \), \( \xi_n = 1/(88n^5) \) when \( n \) is even.

**Proof:** For any squarefree integer \( m \) that is relatively prime to \( N \), let \( \mathcal{W}_m(\Sigma) \) denote the set of elements \( f \in V(\mathbb{Z}) \) such that \( m^2 \mid \Delta(f) \), and such that the reduction of \( f \) modulo \( p^2 \) belongs to \( \Sigma_p \) for every \( p \mid N \). Note that \( \mathcal{W}_m(\Sigma) \) is a union of

\[
\gamma(\Sigma, N, m) := N^{2n+2}m^{2n+2} \prod_{p \mid m} \left( \frac{\# \Sigma_p}{p^{2n+2}} - \alpha_n(\Sigma, p) \right) = \mathcal{O}(N^{2n+2}m^{2n+\epsilon})
\]

translates of \( m^2N^2V(\mathbb{Z}) \). By inclusion-exclusion and Corollary 6.27, we have for any \( M > 0 \),

\[
\begin{align*}
\# \{ f \in V_n(\Sigma) : H(f) < X \text{ and } \Delta(f) \text{ squarefree} \} &= \sum_{(m,N)=1}^{m \leq M} \mu(m) \# \{ f \in \mathcal{W}_m(\Sigma) : H(f) < X \} + \mathcal{O}(\mathcal{E}(X, N, \epsilon)) \\
&= \sum_{(m,N)=1}^{m \leq M} \mu(m) \gamma(\Sigma, N, m) \left( \frac{2X}{N^2m^2} + \mathcal{O}(1) \right)^{n+1} + \mathcal{O}(\mathcal{E}(X, N, \epsilon)) \\
&= \sum_{(m,N)=1}^{m \leq M} \left( (2X)^{n+1} \mu(m) \prod_{p \mid m} \left( \frac{\# \Sigma_p}{p^{2n+2}} - \alpha_n(\Sigma, p) \right) + \mathcal{O}(N^2X^nm^\epsilon + N^{2n+2}m^{2n+\epsilon}) \right) \\
&= \sum_{(m,N)=1}^{m \leq M} \left( (2X)^{n+1} \alpha_n(\Sigma) + \mathcal{O}(\frac{X^{n+1}}{M^{1-\epsilon}} + N^2M^{1+\epsilon}X^N + N^{2n+2}M^{2n+1+\epsilon} + \frac{X^{n+1+\xi_n+\epsilon}}{M^{3n}} + X^{n+1-\eta_n+\epsilon}) \right)
\end{align*}
\]

Recalling that \( \delta_n = 1/2 \) or \( 1/3 \), we may take \( M = X^{3\eta_n+3\xi_n} \) to obtain the first claim in Theorem 7.1. The second claim follows identically. \( \square \)
Taking $N = 1$ in Theorem 7.1 yields Theorem 6. Theorems 1 and 2 are then immediate consequences of Theorem 6.

Next, we prove lower bounds on the number of $S_n$-fields having bounded discriminant. Let $f(x, y) = a_0 x^n + a_1 x^{n-1} y + \cdots + a_n y^n$ be a real binary $n$-ic form with $a_0 \neq 0$ and nonzero discriminant. Let $\theta$ be the image of $x$ in $\mathbb{R}[x]/(f(x, 1))$, and write $R_f$ for the lattice spanned by

\[ 1, \ z_1 = a_0 \theta, \ z_2 = a_0 \theta^2 + a_1 \theta, \ \ldots, \ z_{n-1} = a_0 \theta^{n-1} + \cdots + a_{n-1} \]

in $\mathbb{R}[x]/(f(x, 1))$. Here we identify $\mathbb{R}[x]/(f(x, 1))$ with $\mathbb{R}^n$ via its real and complex embeddings and by identifying $\mathbb{C} = \mathbb{R} \oplus i \mathbb{R}$ with $\mathbb{R}^2$.

We say that $f(x, y)$ is Minkowski-reduced if the basis $\{1, \ z_1, \ldots, \ z_{n-1}\}$ of $R_f$ is Minkowskireduced. We say that $f(x, y)$, or its $\text{SL}_2(\mathbb{Z})$-orbit, is quasi-reduced if there exists $\gamma \in \text{SL}_2(\mathbb{Z})$ such that $\gamma \cdot f$ is Minkowski-reduced. We add the prefix “strongly” if the relevant lattice has a unique Minkowski-reduced basis. The relevance of being strongly quasi-reduced is contained in the following lemma.

Lemma 7.2 Let $n \geq 3$ and let $f(x, y)$ and $f^*(x, y)$ be strongly quasi-reduced integral binary $n$-ic forms. Suppose the corresponding rank-$n$ rings $R_f$ and $R_{f^*}$ are isomorphic. Then $f(x, y)$ and $f^*(x, y)$ are $\text{SL}_2(\mathbb{Z})$-equivalent.

Proof: It suffices to assume $f(x, y) = a_0 x^n + \cdots + a_n y^n$ and $f^*(x, y) = a_0^* x^n + \cdots + a_n^* y^n$ are strongly Minkowski-reduced with $R_f \simeq R_{f^*}$. We show $f(x, y) = f^*(x, y)$. Let $\phi : R_f \to R_{f^*}$ be a ring isomorphism. By the uniqueness of Minkowski-reduced bases, $\phi$ must map the basis elements $1, \ z_1, \ldots, \ z_{n-1}$ for $R_f$ to the corresponding basis elements $1, \ z_1^*, \ldots, \ z_{n-1}^*$ for $R_{f^*}$. Let $\theta$ denote the image of $x$ in $\mathbb{Q}[x]/(f(x, 1))$ and $\theta^*$ the image of $x$ in $\mathbb{Q}[x]/(f^*(x, 1))$. Then $\phi(a_0 \theta) = a_0^* \theta^*$ and

\[ a_0^* \theta^2 + a_1^* \theta^* = \phi(a_0 \theta^2 + a_1 \theta) = (a_0^2/a_0) \theta^2 + (a_1 a_0^*/a_0) \theta^* \]

Since $\theta^*$ and $\theta^* \theta^2$ are linearly independent, we have $a_0 = a_0^*$, $a_1 = a_1^*$, and $\phi(\theta) = \theta^*$, where we extend $\phi$ naturally to $R_f \otimes \mathbb{Q} = \mathbb{Q}[x]/(f(x, 1))$. Then since $\phi(z_{n-1}) = z_{n-1}^*$, we have $a_i = a_i^*$ for $i = 0, \ldots, n-2$. Finally $\phi(-a_{n-1} \theta - a_n) = \phi(\theta z_{n-1}) = \phi(\theta^* z_{n-1}^*) = -a_{n-1}^* \theta^* - a_n^*$. Hence $a_{n-1} = a_{n-1}^*$ and $a_n = a_n^*$.

Proof of Theorem 3: The condition of being strongly quasi-reduced is open in $V_n(\mathbb{R})$. Therefore, given a strongly quasi-reduced element $f \in V_n(\mathbb{R})$, there exists an open neighbourhood $\mathcal{B}$ of $f$ in which every element is strongly quasi-reduced. Moreover, since the action of $\text{SL}_2(\mathbb{Z})$ on $V_n(\mathbb{R})$ is discrete, we may ensure that no two elements of $\mathcal{B}$ are $\text{SL}_2(\mathbb{Z})$-equivalent. We may further scale $\mathcal{B}$ in order to assume that every element in $\mathcal{B}$ has discriminant bounded by 1.

Consider the set $\mathcal{B}_X := X^{1/(2n-2)} \cdot \mathcal{B}$. No two elements in it are $\text{SL}_2(\mathbb{Z})$-equivalent and every element in it is strongly quasi-reduced. Therefore the rings corresponding to any two elements in $\mathcal{B}_X$ are isomorphic. On the other hand, applying Theorem 7.1, we see that $\gg X^{(n+1)/(2n-2)}$ integral elements in $\mathcal{B}_X$ have discriminant less than $X$ and correspond to maximal orders in degree-$n$ number fields. Since these rings are pairwise nonisomorphic, so are their fields of fractions. Hence we have constructed $\gg X^{(n+1)/(2n-2)}$ nonisomorphic degree-$n$ number fields of absolute discriminant less than $X$. Restricting to counting forms that have squarefree discriminant yields $\gg X^{(n+1)/(2n-2)}$ nonisomorphic $S_n$-number fields.

We note that Theorem 7.1 also allows us to construct $\gg X^{(n+1)/(2n-2)} S_n$-number fields satisfying any finite set of splitting conditions.
A Computations of the local densities $\alpha_n(p), \beta_n(p)$

Let $n \geq 2$ be a fixed integer. For a prime $p$, let $\alpha_n(p)$ denote the density of the set of binary $n$-ic forms having discriminant indivisible by $p^2$, and let $\beta_n(p)$ denote the density of binary $n$-ic forms whose associated rank-$n$ rings are maximal at $p$. In this section, we compute $\alpha_n(p)$ and $\beta_n(p)$ for all integers $n \geq 2$ and all primes $p$.

Proposition A.1 We have $\alpha_2(2) = 1/2$ and $\alpha_n(2) = 3/8$ for $n \geq 3$. For odd primes $p$, we have

$$\alpha_n(p) = \begin{cases} 
(1 - \frac{1}{p}) \left(1 + \frac{1}{p} - \frac{1}{p^2}\right) & \text{if } n = 2, \\
(1 - \frac{1}{p})^2 \left(1 + \frac{1}{p}\right)^2 & \text{if } n = 3, \\
(1 - \frac{1}{p})^2 \left(1 + \frac{2}{p} - \frac{2}{p^2} + \frac{1}{p^3}\right) & \text{if } n = 4, \\
(1 - \frac{1}{p})^2 \left(1 + \frac{1}{p}\right) \left(1 + \frac{1}{p} - \frac{1}{p^2}\right) & \text{if } n \geq 5.
\end{cases}$$

Proof: For $j \geq 0$, $n \geq 1$, and $p$ prime, we let $\nu_j(n, p)$ denote the density within monic degree-$n$ integer polynomials of the set of those whose discriminants have $p$-adic valuation $j$. Then $\nu_0(n, p)$ and $\nu_1(n, p)$ are computed in [1, Proposition 6.4 and Theorem 6.8]:

$$\nu_0(n, p) = \begin{cases} 
1 & \text{if } n = 1; \\
1 - p^{-1} & \text{if } n \geq 2.
\end{cases}$$

$$\nu_1(n, p) = \begin{cases} 
0 & \text{if } p = 2 \text{ or } n = 1; \\
p^{-1}(1 - p^{-1}) & \text{if } n = 2, p \neq 2; \\
p^{-1}(1 - p^{-1})^2 & \text{if } n = 3, p \neq 2; \\
(1 - p^{-1})^2(1 - (-p)^{-n})(1 + p)^{-1} & \text{if } n \geq 4, p \neq 2.
\end{cases}$$

To compute the densities $\alpha_n(p)$, we partition the set of integral binary $n$-ic forms $f(x, y) = a_0 x^n + a_1 x^{n-1} y + \cdots + a_n y^n$ whose discriminants are not divisible by $p^2$ into three subsets, and compute each of their densities. For any binary form $f(x, y)$ in $\mathbb{Z}[x, y]$ or in $(\mathbb{Z}/p^2\mathbb{Z})[x, y]$, we write $\tilde{f}(x, y)$ for its reduction modulo $p$.

Subset 1: The set of $f(x, y)$ with $p \nmid a_0$ and $p^2 \nmid \Delta(f)$. Here, for any fixed leading coefficient $a_0 \not\equiv 0 \pmod{p}$, the density of $f(x, y)$ having discriminant indivisible by $p^2$ is simply given by $\nu_0(n, p) + \nu_1(n, p)$. Therefore, the $p$-adic density of this subset is equal to

$$\left(1 - \frac{1}{p}\right)(\nu_0(n, p) + \nu_1(n, p)).$$

Subset 2: The set of $f(x, y)$ with $p | a_0$, $p \nmid a_1$, and $p^2 \nmid \Delta(f)$. In this case, we begin by proving that the density of elements $f$ with fixed $a_0$ and $a_1$ and with $p^2 \nmid \Delta(f)$ is the same as the density of binary $(n - 1)$-ic forms $g$, with fixed leading coefficient $a_1$ such that $p^2 \nmid \Delta(g)$. Indeed, given any $(a_2, \ldots, a_n) \in (\mathbb{Z}/p^2\mathbb{Z})^{n-1}$, we write

$$f_{a_2, \ldots, a_n}(x, y) = a_0 x^n + a_1 x^{n-1} y + a_2 x^{n-2} y^2 + \cdots + a_n y^n \in (\mathbb{Z}/p^2\mathbb{Z})[x, y],$$

$$g_{a_2, \ldots, a_n}(x, y) = a_1 x^{n-1} + a_2 x^{n-2} y + \cdots + a_n y^{n-1} \in (\mathbb{Z}/p^2\mathbb{Z})[x, y].$$
Hence the density of this third subset is
\[ H(x, y)^3 \] for some linear form \( h(x, y) \) has a factor of the form
\[ h(x, y)^3 \] for some linear form \( h(x, y) \) has a factor of the form
\[ j(x, y)^2 \] where \( j \) is a binary form of degree at least 2. Since \( f_{a_0, a_1} (x, y) \equiv y \sum_{a_0, a_1} (x, y) \), and since \( y \) does not divide \( \sum_{a_0, a_1} (x, y) \), we see that \( f_{a_0, a_1} (x, y) \) admits such a factor. Indeed, if \( f_{a_0, a_1} (x, y) \) does. Hence \( S_j^{(1)} = S_j^{(1)} \). On the other hand, we have

\[
\| S_f \| = \# \{ (a_2, \ldots, a_n) \in (\mathbb{Z}/p^2 \mathbb{Z})^{n-1} \setminus S_j^{(1)} : \exists r \in \mathbb{Z}/p\mathbb{Z}, (x - r)^2 \mid f_{a_0, \ldots, a_n} (x, 1), p^2 \mid f_{a_0, \ldots, a_n} (r, 1) \} = \frac{1}{p} \# \{ (a_2, \ldots, a_n) \in (\mathbb{Z}/p^2 \mathbb{Z})^{n-1} \setminus S_j^{(1)} : \exists r \in \mathbb{Z}/p\mathbb{Z}, (x - r)^2 \mid f_{a_0, \ldots, a_n} (x, 1), p \mid f_{a_0, \ldots, a_n} (r, 1) \} = \# \{ (a_2, \ldots, a_n) \in (\mathbb{Z}/p^2 \mathbb{Z})^{n-1} \setminus S_j^{(1)} : \exists r \in \mathbb{Z}/p\mathbb{Z}, (x - r)^2 \mid g_{a_0, \ldots, a_n} (x, 1), p \mid g_{a_0, \ldots, a_n} (r, 1) \} = \# S_j^{(2)}.
\]

The density \( \# S_j^{(1)} \) is \( \nu_0(n - 1, p) + \nu_1(n - 1, p) \). Taking into account that \( p \mid a_0 \) and \( p \nmid a_1 \), we see that the density of this second subset is

\[
\frac{1}{p} \left( 1 - \frac{1}{p} \right) (\nu_0(n - 1, p) + \nu_1(n - 1, p)).
\]

**Subset 3:** The set of \( f(x, y) \) with \( p \mid a_0, p \mid a_1 \), and \( p^2 \nmid \Delta(f) \). Note that we already have \( p \mid \Delta(f) \) in this case. To ensure that \( p^2 \nmid \Delta(f) \), we must have \( p > 2, p^2 \nmid a_0, \) and \( p \nmid a_2 \). Indeed, if \( p = 2 \), then since \( 2 \mid \Delta(f) \), we have \( 4 \mid \Delta(f) \); if \( p^2 \nmid a_0 \), then \( p^2 \) (weakly) divides \( \Delta(f) \); and if \( p \mid a_2 \), then \( y^3 \mid f \) and so \( p^2 \) (strongly) divides \( \Delta(f) \). As polynomials in \( a_0, \ldots, a_n \), we have

\[
\Delta(a_0 x^n + \cdots + a_n) \equiv -4a_0 a_2^3 \Delta(a_2 x^{n-2} + \cdots + a_n) \mod a_0^2 a_0 a_1^2.
\]

Hence if \( p > 2, p^2 \nmid a_0, \) and \( p \nmid a_2 \), then \( p^2 \nmid \Delta(f) \) if and only if \( p \nmid \Delta(a_2 x^{n-2} + a_3 x^{n-3} y + \cdots a_n y^{n-2}) \). Hence the density of this third subset is

\[
\frac{1}{p^2} \left( 1 - \frac{1}{p} \right)^2 \nu_0(n - 2, p).
\]

Adding together these three densities yields the proposition. \( \square \)

Next, we compute the value of \( \beta_n(p) \) for integers \( n \geq 2 \) and primes \( p \).

**Proposition A.2** We have

\[
\beta_n(p) = \begin{cases} 
(1 - \frac{1}{p}) \left( 1 + \frac{1}{p} - \frac{1}{p^3} \right) & \text{if } n = 2; \\
(1 - \frac{1}{p^2}) \left( 1 - \frac{1}{p^3} \right) & \text{if } n \geq 3.
\end{cases}
\]
\textbf{Proof:} The density of monic degree-$n$ integer polynomials that are maximal at $p$ was computed in \cite[Proposition 3.5]{AshBrakenhoffZarrabi2007} to be $1 - p^{-2}$ for all $n \geq 2$ and all primes $p$.

We compute $\beta_n(p)$ by working over $\mathbb{Z}_p$. Fix a binary $n$-ic form $f(x,y) \in V_n(\mathbb{Z}_p)$. Suppose $f(x,y) \mod p$ factors as $y^k g(x,y)$, where $g(x,y)$ is a binary $(n-k)$-ic form over $\mathbb{F}_p$ with nonzero $x^{n-k}$-term for some $k \in \{0, \ldots, n\}$. Then, by Hensel’s lemma, $f(x,y)$ factors as $h_1(x,y)h_2(x,y)$ where $h_1(x,y) \in \mathbb{Z}_p[x]$ is a binary $k$-ic form such that $h_1(x,y) \mod p$ is $y^k$ and $h_2(x,y) \in \mathbb{Z}_p[x]$ is a binary $(n-k)$-ic such that $h_2(x,y) \mod p$ is $g(x,y)$. By scaling $h_1$ and $h_2$, we may further assume that the leading coefficient of $h_2(x,y)$ is $1$.

Since $h_1(x,y)$ and $h_2(x,y)$ share no common factors $\mod p$, the rank-$n$ ring over $\mathbb{Z}_p$ associated to $f(x,y)$ is isomorphic to the product of the rings associated to $h_1(x,y)$ and $h_2(x,y)$. Since $h_1(x,y)$ reduces to a unit times $y^k$ modulo $p$, the rank-$k$ ring associated to $h_1(x,y)$ is always maximal when $k \leq 1$ and is maximal when $k \geq 2$ if and only if $p^2$ does not divide the $x^k$-coefficient. On the other hand, $h_2(x,y)$ is monic, and so the probability that it is maximal is exactly $1 - p^{-2}$ when $n-k \geq 2$, and $1$ when $n-k = 1$. When $k = n$, $f(x,y)$ is a multiple of $p$ and is automatically nonmaximal. Summing over $k$, we have for $n \geq 3$,

$$
\beta_n(p) = \sum_{k=0}^{n-1} \frac{1}{p^k} \left(1 - \frac{1}{p}\right)^{n-k} = \frac{1}{1 - \frac{2}{p^2}} \left(1 - \frac{1}{p^2}\right).$$

When $n = 2$, we have

$$
\beta_2(p) = \left(1 - \frac{1}{p}\right) \left(1 - \frac{1}{p^2}\right) + \frac{1}{p^2} \left(1 - \frac{1}{p}\right)^2 = \left(1 - \frac{1}{p}\right) \left(1 + \frac{1}{p} - \frac{1}{p^2}\right).
$$

This concludes the proof of Proposition A.2. \qed

\textbf{Acknowledgments}

The first-named author was supported by a Simons Investigator Grant and NSF Grant DMS-1001828. The second-named author was supported by an NSERC Discovery Grant and Sloan Research Fellowship. The third-named author was supported by an NSERC Discovery Grant.

\textbf{References}

\begin{enumerate}
\item A. Ash, J. Brakenhoff, and T. Zarrabi, Equality of Polynomial and Field Discriminants, \textit{Experimental. Math.} \textbf{16} (2007), 367–374.
\item M. Bhargava, The density of discriminants of quintic rings and fields, \textit{Ann. of Math. (2)} \textbf{172} (2010), no. 3, 1559–1591.
\item M. Bhargava, The geometric sieve and the density of squarefree values of polynomial discriminants and other invariant polynomials, \url{http://arxiv.org/abs/1402.0031v1}.
\item M. Bhargava, Most hyperelliptic curves over $\mathbb{Q}$ have no rational points, \url{https://arxiv.org/abs/1308.0395v1}.
\item M. Bhargava, Galois groups of random integer polynomials and van der Waerden’s Conjecture, \url{https://arxiv.org/abs/2111.06507v1}.
\end{enumerate}
[6] M. Bhargava, B. Gross, and X. Wang, A positive proportion of locally soluble hyperelliptic curves over \( \mathbb{Q} \) have no point over any odd degree extension, *J. Amer. Math. Soc.* **30** (2017), 451–493.

[7] M. Bhargava, B. Gross, and X. Wang, Arithmetic invariant theory II: Pure inner forms and obstructions to the existence of orbits, *Representations of reductive groups*, Progr. Math., **312**, Birkhäuser/Springer, Cham., 2015, 139–171.

[8] M. Bhargava and W. Ho, On average sizes of Selmer groups and ranks in families of elliptic curves having marked points, preprint.

[9] M. Bhargava and A. Shankar, Binary quartic forms having bounded invariants, and the boundedness of the average rank of elliptic curves, *Ann. of Math.* (2), **181** (2015), no. 1, 191–242.

[10] M. Bhargava, A. Shankar and X. Wang, Squarefree values of polynomial discriminants I, *Invent. Math.* **228** (2022), no. 3, 1037–1073.

[11] B. J. Birch and J. R. Merriman, Finiteness theorems for binary forms with given discriminant, *Proc. London Math. Soc.* (3), **24**, (1972), 385–394.

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

[13] H. Davenport and H. Heilbronn, On the density of discriminants of cubic fields II, *Proc. Roy. Soc. London Ser. A* **322** (1971), no. 1551, 405–420.

[14] T. Ekedahl, An infinite version of the Chinese remainder theorem, *Comment. Math. Univ. St. Paul.* **40** (1991), 53–59.

[15] A. Eskin and Y. Katznelson, Singular symmetric matrices, *Duke Math J.* **79**, No. 2 (1995), 515–547.

[16] I. M. Gelfond, M. M. Kapranov and A. V. Zelevinsky, Discriminants, Resultants and Multidimensional Determinants, Springer Science+Business Media, New York, 1994.

[17] W. Ho, A. Shankar and I. Varma, Odd degree number fields with odd class number, *Duke Math. J.*, **167** (2018), no. 5, 995–1047.

[18] K. S. Kedlaya, A construction of polynomials with squarefree discriminants, *Proc. Amer. Math. Soc.* **140** (2012), 3025–3033.

[19] J. Nakagawa, Binary forms and orders of algebraic number fields, *Invent. Math.* **97** (1989), no. 2, 219–235.

[20] J. Nakagawa, Erratum: “Binary forms and orders of algebraic number fields” [Invent. Math. 97 (1989), no. 2, 219–235], *Invent. Math.* **105** (1991), no. 2, 443.

[21] J. Nakagawa, Binary forms and unramified \( A_n \)-extensions of quadratic fields, *J. Reine Angew. Math.* **406**, (1990), 167–178.

[22] J. Nakagawa, Correction to the paper: “Binary forms and unramified \( A_n \)-extensions of quadratic fields” [J. Reine Angew. Math. 406 (1990), 167–178], *J. Reine Angew. Math.* **413**, (1991), 220.
[23] B. Poonen, Bertini theorems over finite fields, *Ann. of Math.* (2) **160** (2004), no. 3, 1099–1127.

[24] G. C. Sanjaya and X. Wang, On the squarefree values of $a^4 + b^3$. *Math. Ann.* To appear.

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

[26] W. Schmidt, Asymptotic formulae for point lattices of bounded discriminant and subspaces of bounded height, *Duke Math J.* **35** (1968), 327–339.

[27] A. N. Shankar, A. Shankar, and X. Wang, Large families of elliptic curves ordered by conductor, *Compos. Math.* **157** (2021), no. 7, 1538–1583.

[28] A. Shankar and J. Tsimerman, Counting $S_5$-fields with a power saving error term, *Forum Math. Sigma* **2** (2014), e13, 8 pp.

[29] K. Uchida, Unramified extensions of quadratic number fields, II, *Tohoku Math. J.* **22** (1970), 220–224.

[30] M. M. Wood, Rings and ideals parameterized by binary $n$-ic forms, *J. Lond. Math. Soc.* (2), **83**, (2011), no. 1, 208–231.

[31] M. M. Wood, Parametrization of ideal classes in rings associated to binary forms, *J. Reine Angew. Math.* **689** (2014), 169–199.

[32] Y. Yamamoto, On unramified Galois extensions of quadratic number fields, *Osaka J. Math.* **7** (1970), 57–76.

[33] K. Yamamura, On unramified Galois extensions of real quadratic number fields, *Osaka J. Math.* **23** (1986), no. 2, 471–478.