Lower bounds on heights of odd degree points of hyperelliptic curves

Jef Laga and Jack A. Thorne

July 16, 2025

Abstract

We develop a reduction theory for the representation of SL_n on pairs of symmetric $n \times n$ matrices. We apply this theory to the pencils of quadrics arising from divisors on hyperelliptic curves. We use these results to show that, in a density 1 family, an odd degree point P of degree at most $2g-1$ on the hyperelliptic curve $z^2 = f_0x^{2g+2} + f_1x^{2g+1}y + \cdots + f_{2g+2}y^{2g+2}$ cannot have small Weil height.

Contents

1 Introduction

1.1 Results

Let $g \geq 1$ be an integer. Consider the set of integral binary forms $f(x,y) = f_0x^{2g+2} + f_1x^{2g+1}y + \cdots + f_{2g+2}y^{2g+2} \in \mathbb{Z}[x,y]$ of degree $2g+2$ and nonzero discriminant, ordered by the height $\mathrm{Ht}(f) = \max |f_i|$. Such an f determines a hyperelliptic genus-g curve X_f over $\mathbb{Q}$ with equation

$$X_f\colon z^2 = f(x,y) \tag{1.1}$$

in weighted projective space $\mathbb{P}(1,1,g+1)$. When ordered by height, a positive proportion of X_f are everywhere locally soluble, i.e., have $\mathbb{R}$-points and $\mathbb{Q}_p$-points for all primes p [?].

If $g = 1$, Bhargava has shown that among the everywhere locally soluble X_f, a positive proportion have a rational point, and a positive proportion have not [?]. When X_f does have a rational point, it is natural to ask: how complicated must such a point be? In this paper, we show that the heights of such points are, typically, not too small:

Theorem 1.1. *Let $\epsilon > 0$ be arbitrary and $g = 1$. Then for 100% of integral binary quartic forms f (ordered by height), every rational point $P = (x : y : z) \in X_f(\mathbb{Q})$ satisfies*

$$h((x:y)) \geq \left(\frac{5}{4} - \epsilon\right) \log \mathrm{Ht}(f), \tag{1.2}$$

where $h(\alpha)$ denotes the logarithmic Weil height of an element $\alpha \in \mathbb{P}^1(\bar{\mathbb{Q}})$.

We prove a similar theorem for hyperelliptic curves of every genus $g \geq 1$. Bhargava–Gross–Wang have shown that among the everywhere locally soluble X_f, a positive proportion have no points over any odd degree extension of $\mathbb{Q}$ [?]. Analogously to the $g = 1$ case, it seems reasonable to expect (using the heuristics of Poonen and Rains [?]) that a positive proportion of X_f do have an algebraic point of odd degree at most $g + 1$, and again we can ask about the complexity of such a point.

Theorem 1.2. *Let $\epsilon > 0$ be arbitrary. Then for 100% of integral binary forms f of degree $2g+2$ and nonzero discriminant, every algebraic point $P = (x : y : z) \in X_f(\bar{\mathbb{Q}})$ of odd degree $m \leq 2g - 1$ satisfies*

$$m \cdot h((x : y)) \geq \left(1 + \frac{1}{2g+2} - \epsilon\right) \log \mathrm{Ht}(f). \tag{1.3}$$

In fact, we prove a slightly more general result that applies to effective divisors of degree $2g - 1$ on X_f, see Theorem **??**.

Theorems **??** and **??** are analogous to results proved in our previous work [?] for the family of odd hyperelliptic curves of genus g (roughly speaking, the subfamily defined by the conditions $f_0 = 0$, $f_1 = 1$). We compare the methods used in this work to those of [?] below.

1.2 Methods

Let $n = 2g + 2$ and consider the representation of SL_n on the space V of pairs of symmetric $n \times n$ matrices, where the action is defined by $g \cdot (A, B) = (g^{-t}Ag^{-1}, g^{-t}Bg^{-1})$. To such a pair (A, B), we may associate an invariant binary form

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

whose coefficients freely generate the ring of polynomial invariants for the SL_n-action on V [?].

A key step in our proof of Theorem **??** is the development of a reduction theory for the $\mathrm{SL}_n(\mathbb{Z})$-action on $V(\mathbb{Z})$. Generally speaking, given an action of an arithmetic group on a set, reduction theory aims to find 'reduced' representatives of orbits for the action, e.g., representatives whose coefficients are small. In our setting, we define for every pair $(A, B) \in V(\mathbb{R})$ with $\mathrm{disc}(f_{A,B}) \neq 0$ an inner product $H_{A,B}$ on $\mathbb{R}^n$, whose formation is $\mathrm{SL}_n(\mathbb{R})$-equivariant. We say that a pair (A, B) is Minkowski (or LLL) reduced if the standard basis of $\mathbb{R}^n$ is Minkowski (or LLL) reduced with respect to $H_{A,B}$. The coefficients of reduced pairs (in either sense) can be absolutely bounded in terms of the coefficients of $f_{A,B}$, and so the problem of reducing elements in $V(\mathbb{Z})$ becomes a problem in the reduction theory of lattices, which is well understood.

Let $X = \mathrm{SL}_n(\mathbb{R})/\mathrm{SO}_n(\mathbb{R})$ denote the symmetric space of inner products on $\mathbb{R}^n$ up to scaling. We call the association $(A, B) \mapsto H_{A,B}$ or, by passage to the quotient, the association

$$\mathcal{R}\colon\ \mathrm{SL}_n(\mathbb{Z})\backslash V(\mathbb{Z})^{\Delta\neq 0} \to \mathrm{SL}_n(\mathbb{Z})\backslash X \tag{1.4}$$

the 'reduction covariant' of the representation V, akin to the reduction covariant defined in [?, ?, ?]. In the context of [?], the reduction covariant is used in the algorithmic study of n-descent on elliptic curves, and we similarly expect our reduction covariant to be useful for computational purposes.

In this paper, we instead employ the reduction covariant for theoretical purposes, and relate it to the descent theory of hyperelliptic curves. Suppose $f(x, y) \in \mathbb{Z}[x, y]$ is a binary form of degree n and nonzero discriminant, and let V_f be the subset of (A, B) with $f_{A,B} = f$. Given an effective $\mathbb{Q}$-rational divisor D on X_f of degree $2g - 1$, we build on the methods of [?] to produce an integral orbit $\mathrm{SL}_n(\mathbb{Z}) \cdot (A, B) \subset V_f(\mathbb{Z})$, which then has an associated reduction covariant $[H_{A,B}] \in \mathrm{SL}_n(\mathbb{Z})\backslash X$. We show that (in a precise sense) if D has small height, then $\mathbb{Z}^n$ has a primitive element of small norm with respect to $H_{A,B}$. On the other hand, we show that the reduction covariants of orbits of height bounded by X are equidistributed as $X \to \infty$, with respect to the natural measure on the target of (**??**). Since the subset of inner products in $\mathrm{SL}_n(\mathbb{Z})\backslash X$ admitting a nonzero vector of small norm has small measure, we conclude that only a small proportion of curves X_f can admit a divisor D of small height.

The construction of integral orbits alluded to in the previous paragraph is a key technical innovation in this paper, which we achieve by introducing a new method to construct such orbits. This is essentially a local question, so we may begin with a rational odd degree divisor over $\mathbb{Q}_p$ and construct an element of $V(\mathbb{Z}_p)$. Such a construction was given in our setting in [**?**, Theorem 15], although their argument contains gaps, which were corrected by Swaminathan [**?**, Appendix 4A]. A construction in the setting of odd hyperelliptic curves was given by Bhargava–Gross [**?**, Propostion 19], but this includes a reduction step that makes it not useful for our purpose here (where we have to be able to show that a specific vector is integral with respect to the given integral structure, in order to know that its length cannot be too small too often). In our previous work [**?**], we resolved this by a more careful construction that still, like the one in [**?**], relies on an induction of the degree of the divisor.

The construction of integral orbits we give here uses the interpretation of these orbits given by Wood [**?**], namely as corresponding to certain classes of ideals for the ring associated to the binary form $f(x, y)$. This ring may be thought of as the global sections of the structure sheaf on the subscheme S_f of $\mathbb{P}^1_{\mathbb{Z}_p}$ cut out by f. We associate to a divisor D a torsion-free sheaf $\mathcal{O}_X(D)$ on the hyperelliptic curve X_f over $\mathbb{Z}_p$ defined by the same equation $z^2 = f(x, y)$; this curve contains S_f as the locus $z = 0$, and the integral orbit we construct is the one associated to the ideal class of $H^0(S_f, \mathcal{O}_X(D)|_{S_f})$. The hard part is to show that this ideal class has the right properties to correspond to an integral orbit; this we accomplish using coherent Grothendieck duality for the scheme S_f.

We now describe the relation between the methods employed here and those used to prove the main theorems of our previous work on odd hyperelliptic curves [**?**]. In that paper, we employed a broadly similar strategy, using the equidistribution of reduction covariants, to obtain height lower bounds in a density-1 family of monic odd genus-g hyperelliptic curves, using the representation of SO_{2g+1} on self-adjoint traceless $(2g+1) \times (2g+1)$ matrices. However, there are some key differences. First, the representation of SL_n considered in this paper falls outside the scope of Vinberg theory, so the general construction of the reduction covariant of [**?**] does not apply here. In fact, for the representation of SL_n there are infinitely many reduction covariants, and it is important to select one with the correct properties. Second, the reduction covariant here equidistributes over all integral orbits of nonzero discriminant in V, while in [**?**] it only equidistributes over the *irreducible* integral orbits, and handling the reducible orbits required an additional argument using the methods of [**?**]. Finally, the construction of integral orbits described above uses a different method to the one employed in [**?**]. It seems likely that the method used here would also work in that setting.

1.3 Organization

In §**??**, we recall some results of Wood [**?**, **?**] concerning binary forms and the rank n rings they define. We emphasize the role played by the closed subscheme $S_f \subset \mathbb{P}^1_{\mathbb{Z}}$ cut out by the form f and show, using coherent duality, how certain sheaves on $\mathcal{M}$ give rise to integral orbits for the representation V. In §**??** we describe the representation V of SL_n together with its rational and integral orbits. We also introduce the reduction covariant of an element $(A, B) \in V(\mathbb{R})$ of nonzero discriminant. In §**??**, we present the crucial construction of rational and integral orbits associated to effective divisors of odd degree on X_f satisfying some conditions. Moreover, we compute the norm of a certain element in $\mathbb{Z}^n$ with respect to the reduction covariant of this integral orbit. In §**??**, we prove that the reduction covariant (**??**) equidistributes, by adapting the geometry-of-numbers methods of Bhargava [**?**]. In §**??**, we combine all of these ingredients to prove the main theorems of the introduction.

2 Binary forms and ideal classes

There is a long-studied association between binary forms of fixed degree n and rings of rank n (see e.g. [?]), which has been analysed in detail by Wood [?], who also studied the relation between ideal classes for these rings and orbits in the representation V of SL_n studied later in this paper [?]. In this section, we first recall some of Wood's results, and complement them with a geometric construction of orbits associated to sheaves on the zero locus of S_f. Our desire to include forms that are not primitive makes various proofs more technical.

Let $n \geq 2$ be an integer, and let A be a Dedekind domain of fraction field K of characteristic 0. If $k \in \mathbb{Z}$, write $\mathcal{O}_{\mathbb{P}^1_A}(k)$ for the usual sheaf on $\mathbb{P}^1_A$ whose global sections are the forms in $A[x,y]$ that are homogeneous of degree k. Let

$$f(x,y) = f_0 x^n + f_1 x^{n-1} y + \cdots + f_n y^n \in H^0(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(n))$$

be a form whose discriminant $\Delta(f) \in A$ is nonzero, define $S_f \subset \mathbb{P}^1_A$ to be the closed subscheme defined by the vanishing of f, and $R_f = H^0(S_f, \mathcal{O}_{S_f})$.

Proposition 2.1. *The ring R_f is a free A-algebra of rank n.*

Proof. We consider the short exact sequence of coherent sheaves on $\mathbb{P}^1_A$:

$$0 \to \mathcal{O}_{\mathbb{P}^1_A}(-n) \overset{\times f}{\to} \mathcal{O}_{\mathbb{P}^1_A} \to \mathcal{O}_{S_f} \to 0.$$

After passing to global sections, we get a short exact sequence

$$0 \to A \to R_f \to H^1(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(-n)) \to 0.$$

Since $H^1(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(-n))$ is a free A-module of rank $n-1$, this completes the proof. □

Proposition 2.2. *If $f(x,y)$ is primitive, then the scheme S_f is affine.*

To say that $f(x,y)$ is primitive is to say that its content (i.e. the ideal of A generated by $f_0, \ldots, f_n$) is the unit ideal.

Proof. It suffices to prove this affine locally on $\operatorname{Spec} A$, so we can assume that there exists a homogeneous form $g(x,y) \in A[x,y]$ of some degree m such that for every non-zero prime ideal $P \leq A$, the reductions modulo P $\overline{f}, \overline{g} \in (A/P)[x,y]$ are non-zero and cut out disjoint closed subschemes of $\mathbb{P}^1_{A/P}$. Then S_f is identified with a closed subscheme of the distinguished open $D^+(g) \subset \mathbb{P}^1_A$, and is therefore affine. □

If $k \in \mathbb{Z}$, we write $\mathcal{O}_{S_f}(k)$ for the pullback of $\mathcal{O}_{\mathbb{P}^1_A}(k)$ to S_f. We define $I_f = H^0(S_f, \mathcal{O}_{S_f}(n-3))$ and $J_f = H^0(S_f, \mathcal{O}_{S_f}(n-2))$. Then I_f, J_f are finite R_f-modules that are free over A of rank n (the proof is a twist of the proof of Proposition ??). Taking cohomology, we see that there are short exact sequences

$$0 \to H^0(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(n-3)) \to I_f \to H^1(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(-3)) \to 0 \tag{2.1}$$

and

$$0 \to H^0(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(n-2)) \to J_f \to H^1(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(-2)) \to 0. \tag{2.2}$$

The group $H^1(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(-2))$ is free of rank 1 over A, a basis being given by the class represented in Cech cohomology with respect to the cover $D^+(x), D^+(y)$ by the element $(xy)^{-1} \in H^0(D^+(xy), \mathcal{O}_{\mathbb{P}^1_A}(-2))$. We write $\zeta : J_f \to A$ for the A-linear form induced by the exact sequence (**??**) and this choice of basis element.

If $a(x,y) \in H^0(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(1))$ is a linear form, write $\mathrm{ev}_a : I_f \to A$ for the composite

$$\mathrm{ev}_a \colon I_f \to H^1(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(-3)) \to H^1(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(-2)) \to A,$$

where the second map is induced by multiplication by a, and the third is ζ.

Lemma 2.3. *For every linear form a, ev_a equals the composite*

$$\beta_a \colon I_f \to J_f \to A,$$

where the first map is induced by passage to global sections of the morphism $\times a : \mathcal{O}_{S_f}(n-3) \to \mathcal{O}_{S_f}(n-2)$, and the second map is ζ.

Proof. This follows by considering the commutative diagram with exact rows

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{O}_{\mathbb{P}^1_A}(-3) \arrow[r] \arrow[d] & \mathcal{O}_{\mathbb{P}^1_A}(n-3) \arrow[r] \arrow[d] & \mathcal{O}_{S_f}(n-3) \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathcal{O}_{\mathbb{P}^1_A}(-2) \arrow[r] & \mathcal{O}_{\mathbb{P}^1_A}(n-2) \arrow[r] & \mathcal{O}_{S_f}(n-2) \arrow[r] & 0,
\end{tikzcd}$$

where the vertical maps are all multiplication by a. □

Proposition 2.4. *Let $\mathcal{M}$ be a coherent sheaf on S_f such that $H^1(S_f, \mathcal{M}) = 0$ (this is automatic if f is primitive), and let $M = H^0(S_f, \mathcal{M})$. Then there is a canonical isomorphism*

$$\mathrm{Hom}_{\mathcal{O}_{S_f}}(\mathcal{M}, \mathcal{O}_{S_f}(n-2)) \cong \mathrm{Hom}_A(M, A),$$

given by passage to global sections and then composition with ζ.

Proof. We will use coherent Grothendieck duality in the form of [**?**, **?**]. For a Noetherian scheme X, denote by $\mathbf{D}(X)$ the derived category of $\mathcal{O}_X$-modules and $\mathbf{D}^+_c(X)$ (respectively $\mathbf{D}^-_c(X)$) the full subcategory consisting of complexes whose cohomology sheaves are coherent and vanish in sufficiently negative degrees (respectively positive degrees). The theory associates, to any morphism $\varphi \colon X \to Y$ of finite type schemes over A, a functor $\varphi^! \colon \mathbf{D}^+_c(Y) \to \mathbf{D}^+_c(X)$ satisfying various properties explained in [**?**, Section 3.3], and when φ is furthermore proper, a trace map $\mathrm{tr}_\varphi \colon R\varphi_* \circ \varphi^! \Rightarrow \mathrm{Id}$, which is a natural transformation of endofunctors on $\mathbf{D}^+_c(Y)$ [**?**, §3.4]. Grothendieck–Serre duality then states that for all $F \in \mathbf{D}^-_c(X)$ and $G \in \mathbf{D}^+_c(Y)$, the composition

$$\mathrm{Hom}_{\mathbf{D}(X)}(F, \varphi^! G) \xrightarrow{R\varphi_*} \mathrm{Hom}_{\mathbf{D}(Y)}(R\varphi_* F, R\varphi_* \varphi^! G) \xrightarrow{\mathrm{tr}_\varphi} \mathrm{Hom}_{\mathbf{D}(Y)}(R\varphi_* F, G) \tag{2.3}$$

is an isomorphism of A-modules; this follows from applying $H^0 \circ R\Gamma$ to [**?**, Theorem 3.4.4].

In our concrete situation, we have proper maps $i \colon S_f \hookrightarrow \mathbb{P}^1_A$, $p \colon \mathbb{P}^1_A \to \mathrm{Spec}(A)$ and $\pi \colon S_f \to \mathrm{Spec}(A)$ satisfying $\pi = p \circ i$. We will show that Grothendieck duality for the triple $(\varphi, F, G) = (\pi, \mathcal{M}, A)$ implies the claim of the proposition. In our arguments we will sometimes identify quasi-coherent modules on $\mathrm{Spec}(A)$ with A-modules. We first make the theory explicit for i and p.

- If $\mathcal{F}$ is a coherent sheaf on $\mathbb{P}^1_A$, then $Ri_* i^! \mathcal{F} = R\mathcal{H}om_{\mathbf{D}(\mathbb{P}^1_A)}(i_* \mathcal{O}_{S_f}, \mathcal{F})$ and $\mathrm{tr}_i \colon R\mathcal{H}om_{\mathbf{D}(\mathbb{P}^1_A)}(i_* \mathcal{O}_{S_f}, \mathcal{F}) \to \mathcal{F}$ is given by precomposing with the quotient map $\mathcal{O}_{\mathbb{P}^1_A} \to i_* \mathcal{O}_{S_f}$ and applying the canonical isomorphism $R\mathcal{H}om_{\mathbf{D}(\mathbb{P}^1_A)}(\mathcal{O}_{\mathbb{P}^1_A}, \mathcal{F}) \simeq \mathcal{F}$; see [**?**, Lemma 3.4.3(2) and page 31]. Since S_f is a Cartier divisor, $i_* \mathcal{O}_{S_f}$ is computed by the complex of locally free sheaves

$$[\mathcal{O}_{\mathbb{P}^1_A}(-n) \xrightarrow{\times f} \mathcal{O}_{\mathbb{P}^1_A}]$$

(with $\mathcal{O}_{\mathbb{P}^1_A}$ placed in degree 0), so $R\mathcal{H}om_{\mathbf{D}(\mathbb{P}^1_A)}(i_* \mathcal{O}_{S_f}, \mathcal{F})$ is represented by the complex

$$[\mathcal{F} \xrightarrow{\times f} \mathcal{F}(n)]$$

living in degrees 0 and 1, and tr_i is induced by the map of complexes

$$[\mathcal{F} \to \mathcal{F}(n)] \to [\mathcal{F} \to 0]$$

which is the identity in degree 0 and the zero map in degree 1. If $\mathcal{F}$ is furthermore locally free, the map $\mathcal{F} \xrightarrow{\times f} \mathcal{F}(n)$ is injective, so there is a quasi-isomorphism

$$[\mathcal{F} \to \mathcal{F}(n)] \xrightarrow{\sim} i_* i^* (\mathcal{F}(n))[-1].$$

- If $\mathcal{F}$ is a coherent sheaf on $\mathrm{Spec}(A)$, then $p^!(\mathcal{F}) = \Omega^1_{\mathbb{P}^1_A/A}[1] \otimes_{\mathcal{O}_{\mathbb{P}^1_A}} f^* \mathcal{F}$ by [**?**, Lemma 3.4.3(3)]. Moreover if $\mathcal{F} = A$ then $\mathrm{tr}_p \colon Rp_* p^! A \to A$ equals, when viewed as a map of A-modules $H^1(\mathbb{P}^1_A, \Omega^1_{\mathbb{P}^1_A/A}) \to A$, the trace isomorphism from classical Serre duality fixed in [**?**, Section 2.3].

Using these examples, the isomorphism $\Omega^1_{\mathbb{P}^1_A/A} \cong \mathcal{O}_{\mathbb{P}^1_A}(-2)$, and the natural isomorphism $\pi^! \cong i^! \circ p^!$ [**?**, Equation (3.3.14)], we compute

$$\pi^! A \cong i^! p^! A \cong \mathcal{O}_{S_f}(n-2),$$

in $\mathbf{D}^+_c(S_f)$, and

$$R\iota_* \pi^! A \cong i_* \mathcal{O}_{S_f}(n-2) \cong [\mathcal{O}_{\mathbb{P}^1_A}(-2) \overset{\times f}{\to} \mathcal{O}_{\mathbb{P}^1_A}(n-2)].$$

in $\mathbf{D}^+_c(\mathbb{P}^1_A)$. We observe that the short exact sequence (**??**) defining the map ζ expresses the filtration associated to the quasi-isomorphism $\mathcal{O}_{S_f}(n-2) \cong [\mathcal{O}_{\mathbb{P}^1_A}(-2) \overset{\times f}{\to} \mathcal{O}_{\mathbb{P}^1_A}(n-2)]$, the map ζ itself corresponding to the composite of the Serre duality isomorphism $H^1(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(-2)) \cong A$ with the map $H^0(S_f, \mathcal{O}_{S_f}(n-2)) \to H^1(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(-2))$ determined by the composite

$$i_* \mathcal{O}_{S_f}(n-2) \cong [\mathcal{O}_{\mathbb{P}^1_A}(-2) \overset{\times f}{\to} \mathcal{O}_{\mathbb{P}^1_A}(n-2)] \to \mathcal{O}_{\mathbb{P}^1_A}(-2)[1]$$

in $\mathbf{D}^+_c(\mathbb{P}^1_A)$.

We now connect this to trace. Using the examples above, we see that $\mathrm{tr}_i(p^! A) : Ri_* i^! p^! A \to p^! A$ is the morphism

$$[\mathcal{O}_{\mathbb{P}^1_A}(-2) \overset{\times f}{\to} \mathcal{O}_{\mathbb{P}^1_A}(n-2)] \to \mathcal{O}_{\mathbb{P}^1_A}(-2)[1],$$

showing that $Rp_*(\mathrm{tr}_i(p^! A)) : Rp_* Ri_* i^! p^! A \to Rp_* p^! A$ is the induced morphism

$$J_f = H^0(\mathbb{P}^1_A, i_* \mathcal{O}_{S_f}(n-2)) = H^0(\mathbb{P}^1_A, [\mathcal{O}_{\mathbb{P}^1_A}(-2) \overset{\times f}{\to} \mathcal{O}_{\mathbb{P}^1_A}(n-2)]) \to H^1(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(-2)).$$

On the other hand, by [**?**, Lemma 3.4.3(1)], we have $\mathrm{tr}_\pi = \mathrm{tr}_p \circ Rp_*(\mathrm{tr}_i)$, so $\mathrm{tr}_\pi(A) : R\pi_* \pi^! A \to A$ is the morphism $J_f \to A$ obtained by composing this induced morphism with the Serre duality isomorphism $\mathrm{tr}_p(A) : H^1(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(-2)) \to A$. We have shown that $\mathrm{tr}_\pi(A) = \zeta$. (We note that the references [**?**, **?**]

contain, at points, different choices, that may affect the sign of tr_π, and therefore the truth of the statement $\mathrm{tr}_\pi(A) = \zeta$; however, the truth of Proposition **??**, namely that a certain map of A-modules induced by ζ is an isomorphism, is insensitive to multiplying ζ by a sign, so we do not need to worry about this here.)

We now show how the desired statement follows from (**??**). By assumption, we have $H^1(S_f, \mathcal{M}) = 0$, hence $R\pi_* \mathcal{M} \cong M$. We have shown that there is an isomorphism $\pi^! A \cong \mathcal{O}_{S_f}(n-2)$. We find that ζ induces an isomorphism

$$\mathrm{Hom}_{\mathbf{D}(S_f)}(\mathcal{M}, \mathcal{O}_{S_f}(n-2)) \cong \mathrm{Hom}_{\mathbf{D}(A)}(M, A).$$

Since the embedding of the category of $\mathcal{O}_X$-modules in $\mathbf{D}(X)$ is fully faithful, this is exactly what we needed to show. □

We can use a similar point of view to construct pencils of quadrics over A from suitable sheaves on S_f. Suppose given a coherent sheaf $\mathcal{M}$ on S_f satisfying the following conditions:

- $\mathcal{M}_\eta$ is free of rank 1 at each generic point $\eta \in S_f$.
- There is an isomorphism $\mathcal{M} \cong \mathcal{H}om_{\mathcal{O}_{S_f}}(\mathcal{M}, \mathcal{O}_{S_f}(n-3))$.
- $H^1(S_f, \mathcal{M}) = 0$.

Proposition 2.5. *With assumptions as above, let $M = H^0(S_f, \mathcal{M})$. Then:*

1. *M is a finite R_f-module, A-torsion-free, and $M \otimes_A K$ is free of rank 1 over $R_f \otimes_A K$.*
2. *By passage to global sections, we obtain an A-linear map $M \times M \to I_f \to H^0(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1}(1))^\vee$. Fix an A-basis of M, and let A_x, A_y denote the matrices of the two symmetric bilinear forms $M \times M \to A$ given by composition with ev_x and ev_y, respectively. Then there exists a unit $u \in A^\times$ such that $\det(xA_y - yA_x) = uf(x, y)$.*

Proof. We take each point in turn. Since S_f is projective, M is a finite R_f-module. Since $S_{f,K}$ is a finite étale K-scheme, and $\mathcal{M}_K$ is assumed locally free of rank 1, $M \otimes_A K$ is free of rank 1 over $R_f \otimes_A K$. We must check that M is A-torsion-free. We can check this locally on A, so can assume that A is a DVR with uniformizer π. We are also free to make an étale localisation on A, so can assume that there is a primitive homogeneous linear polynomial $a(x, y) = x - sy \in A[x, y]$ such that $f(s, 1) \neq 0$. Consider the short exact sequence of sheaves on $\mathbb{P}^1_A$

$$0 \to \mathcal{O}_{\mathbb{P}^1_A}(n-3) \overset{\times a}{\to} \mathcal{O}_{\mathbb{P}^1_A}(n-2) \to \mathcal{O}_a \to 0,$$

where $\mathcal{O}_a$ (viewed as a sheaf on $\mathbb{P}^1_A$) is the structure sheaf of the closed subscheme $V(a) \subset \mathbb{P}^1_A$, isomorphic to $\mathrm{Spec}\, A$. This short exact sequence remains exact after pullback along $i : S_f \to \mathbb{P}^1_A$ (because $A[x]/a(x, 1)$ has no $f(x, 1)$-torsion). After making this pullback and applying the functor $\mathrm{Hom}_{\mathcal{O}_{S_f}}(\mathcal{M}, -)$, we get an exact sequence

$$0 \to \mathrm{Hom}_{\mathcal{O}_{S_f}}(\mathcal{M}, \mathcal{O}_{S_f}(n-3)) \to \mathrm{Hom}_{\mathcal{O}_{S_f}}(\mathcal{M}, \mathcal{O}_{S_f}(n-2)) \to \mathrm{Hom}_{\mathcal{O}_{S_f}}(\mathcal{M}, i^* \mathcal{O}_a). \tag{2.4}$$

Applying our hypotheses on $\mathcal{M}$ and Proposition **??**, we see in particular that we have an embedding $M \to \mathrm{Hom}_A(M, A)$. The module $\mathrm{Hom}_A(M, A)$ is certainly A-torsion-free, so this shows that M is itself A-torsion-free.

To prove the second part, we proceed in stages. Let $h(x, y) = \det(xA_y - yA_x) \in A[x, y]$. We will first show that $h(x, y)$ and $f(x, y)$ cut out the same locus in $\mathbb{P}^1_K$, using a calculation over the algebraic closure. This implies that $f(x, y)$ and $h(x, y)$ are equal up to multiple by some element of $K^\times$. To show equality, we will then be free to localise on A and assume that A is a DVR of uniformizer π. We will then show the existence of a pair $(r, s) \in A^2$ such that $h(r, s)$ and $f(r, s)$ have the same (finite) π-adic valuation, forcing $f(x, y)$ and $h(x, y)$ to be equal up to multiplication by units.

It follows from Lemma **??** that if $a(x, y) = rx + sy$, we write $\varphi : M \times M \to I_f$ for the R_f-bilinear pairing, and A_a for the matrix of the symmetric A-bilinear form $\mathrm{ev}_a \circ \varphi : M \times M \to A$, then $A_a = rA_x + sA_y$, and hence $\det(A_a) = \det(sA_y + rA_x) = h(s, -r)$.

We now show that $f(x, y)$ divides $h(x, y)$. This can be checked after base extension to the algebraic closure of K. Suppose that $a(x, y)$ a linear form dividing $f(x, y)$. Then multiplication by a sends I_f to a codimension 1 subspace of J_f. Using Lemma **??**, we see that $\mathrm{ev}_a \circ \varphi = \zeta \circ \beta_a \circ \varphi$ does not have full rank, showing that $\det(A_a) = 0$ and hence $a(x, y)$ divides $h(x, y)$. Since $f(x, y)$ has non-zero discriminant and a was an arbitrary linear factor, this shows that f divides h; since they have the same degree, they are in fact equal up to an element of $K^\times$.

Now we return to the case of general A, and must show that $f, h \in A[x, y]$ are in fact equal up to multiplication by elements of $A^\times$. After localisation, we can assume that A is a DVR of uniformizer π and with infinite residue field k. Let us write $f(x, y) = \pi^m g(x, y)$ where $m \geq 0$ and $g(x, y)$ is primitive. We first treat the case where $m = 0$, i.e. $f(x, y)$ is itself primitive. In this case, we can choose a primitive linear form $a(x, y) = rx + sy \in A[x, y]$ such that $\overline{a}$ does not divide $\overline{f}$ in $k[x, y]$ (where overline denotes reduction modulo π); equivalently, $f(s, -r) \in A^\times$. Forming the exact sequence (**??**), we get a short exact sequence

$$0 \to M \to \mathrm{Hom}_A(M, A) \to \mathrm{Hom}_{\mathcal{O}_{S_f}}(\mathcal{M}, i^*\mathcal{O}_a)$$

The group $\mathrm{Hom}_{\mathcal{O}_{S_f}}(\mathcal{M}, i^*\mathcal{O}_a)$ is in fact 0, because our choice of a means that $i^*\mathcal{O}_a = 0$ (in other words, the section $a \in H^0(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(1))$ is non-vanishing on S_f). In particular, $h(s, -r) \in A^\times$ and we're done in this case.

We now suppose that $m > 0$. In this case, S_f contains $\mathbb{P}^1_k$ as a closed subscheme and the sheaf $i^*\mathcal{O}_a$ is non-zero. Since we assume that $\mathcal{M}_\eta$ is free for each generic point η of S_f, there exists an open subscheme U of S_f, supported over $\operatorname{Spec} k$, and over which $\mathcal{M}$ is free of rank 1. Let us choose a so that $g(s, -r) \in A^\times$ and such that the support of $i^*\mathcal{O}_a$ is contained in U. In this case, we again form the exact sequence (**??**), to obtain

$$0 \to M \to \mathrm{Hom}_A(M, A) \to \mathrm{Hom}_{\mathcal{O}_{S_f}}(\mathcal{M}, i^*\mathcal{O}_a) \to 0.$$

To justify the exactness on the right, it is enough to show that $\mathrm{Ext}^1_{\mathcal{O}_{S_f}}(\mathcal{M}, \mathcal{O}_{S_f}(n-3)) = 0$. The sheaf $\mathcal{E}xt^1_{\mathcal{O}_{S_f}}(\mathcal{M}, \mathcal{O}_{S_f}(n-3))$ vanishes, by [**?**, Proposition 1.6] (note that $\mathcal{M}$ is reflexive because there is an isomorphism $\mathcal{M} \cong \mathcal{H}om_{\mathcal{O}_{S_f}}(\mathcal{M}, \mathcal{O}_{S_f}(n-3))$, so we can apply [**?**, Corollary 1.8]). Therefore

$$\mathrm{Ext}^1_{\mathcal{O}_{S_f}}(\mathcal{M}, \mathcal{O}_{S_f}(n-3)) \cong H^1(S_f, \mathcal{H}om_{\mathcal{O}_{S_f}}(\mathcal{M}, \mathcal{O}_{S_f}(n-3))) \cong H^1(S_f, \mathcal{M}),$$

and this group vanishes, by assumption.

By construction, $\mathrm{Hom}_{\mathcal{O}_{S_f}}(\mathcal{M}, i^*\mathcal{O}_a)$ is isomorphic as A-module to the stalk of $i^*\mathcal{O}_a$ at the point in $\mathbb{P}^1_k$ where $\overline{a}$ vanishes; in other words, to $A/(\pi^m)$. By Lemma **??**, the morphism $M \to \mathrm{Hom}_A(M, A)$ in the above exact sequence is the one arising from $\mathrm{ev}_a \circ \varphi$, so we have shown that $\det(A_a) \in \pi^m A^\times$. Since $f(s, -r) \in \pi^m A^\times$, by construction, this completes the proof. □

3 The action of SL_n on pairs of symmetric matrices

3.1 Basic definitions

Let $n \geq 2$ be an integer and let W be the free $\mathbb{Z}$-module of rank n with basis $e_1, \dots, e_n$. Let $\mathrm{Sym}_2(W^\vee)$ be the free $\mathbb{Z}$-module of symmetric bilinear forms $(-,-)\colon W \times W \to \mathbb{Z}$. Given such a form $(-,-)$, its Gram matrix with respect to the basis $\{e_i\}$ is the matrix $A = ((e_i, e_j))_{1 \leq i,j \leq n}$. We use this to view $\mathrm{Sym}_2(W^\vee)$ as the $\mathbb{Z}$-module of symmetric matrices $A \in \mathrm{Sym}_2(\mathbb{Z}^n) \subset \mathrm{Mat}_n(\mathbb{Z})$. Conversely, given a symmetric matrix $A \in \mathrm{Sym}_2(\mathbb{Z}^n)$, write $(-,-)_A$ for the corresponding symmetric bilinear form on W.

The group $\mathrm{GL}(W)$ and its subgroup $\mathrm{SL}(W)$ act on W and hence by functoriality on $\mathrm{Sym}_2(W^\vee) = \mathrm{Sym}_2(\mathbb{Z}^n)$. This action has the property that $(v, w)_{g \cdot A} = (g^{-1}v, g^{-1}w)_A$ for all $v, w \in V$, $A \in \mathrm{Sym}_2(W^\vee)$ and $g \in \mathrm{GL}(W)$. In terms of matrices, we have the formula $g \cdot A = g^{-t}Ag^{-1}$. The group $\mathrm{SL}(W)$ acts on V via $g \cdot (A, B) = (g^{-t}Ag^{-1}, g^{-t}Bg^{-1})$. The same formulae define actions of the group $\mathrm{SL}_n(R)$ on the R-modules $\mathrm{Sym}_2((W \otimes_\mathbb{Z} R)^\vee) = \mathrm{Sym}_2(R)$ and $V(R) = \mathrm{Sym}_2((W \otimes_\mathbb{Z} R)^\vee) \oplus \mathrm{Sym}_2((W \otimes_\mathbb{Z} R)^\vee)$, for any ring R. (Our formula for the $\mathrm{SL}(W)$-action on V differs from [**?**, **?**], where they consider the action defined by $g \cdot (A, B) = (gAg^t, gBg^t)$, which coincides with our action after replacing g by its inverse transpose. Since the orbits over any ring are the same for both actions, this difference is harmless.)

If R is a ring and $A \in \mathrm{Sym}_2(R^n)$, its discriminant is by definition $\mathrm{disc}(A) = (-1)^{n(n-1)/2}\det(A)$. We define the invariant binary form of $(A, B) \in V(R)$ to be

$$f_{A,B}(x, y) = \mathrm{disc}(Ax - By) = f_0x^n + f_1x^{n-1}y + \cdots + f_ny^n \in R[x, y].$$

The coefficients of the invariant binary form are invariant polynomials of degree n in the coefficients of (A, B) for the $\mathrm{SL}_n(R)$-action, and freely generate the full ring of invariants when $R = \mathbb{Z}$. If $f(x, y) \in R[x, y]$ is a binary form of degree n, write $V_f(R)$ for the $\mathrm{SL}_n(R)$-stable subset of $(A, B) \in V(R)$ satisfying $f_{A,B} = f$.

The discriminant $\Delta(f)$ of a binary form $f = f_0x^n + \cdots + f_ny^n$ is a homogeneous polynomial of degree $2n - 2$ in $f_0, \dots, f_n$. It is uniquely characterized by the property that if $R = K$ is an algebraically closed field and $f = \prod_{i=1}^n(\alpha_ix - \beta_iy)$ for some $\alpha_i, \beta_i \in K$, then

$$\Delta(f) = \prod_{i<j}(\alpha_i\beta_j - \alpha_j\beta_i)^2.$$

If $(A, B) \in V(R)$ we set $\Delta(A, B) = \Delta(f_{A,B})$.

3.2 Description of rational orbits

Let K be a field of characteristic zero and $f(x, y) = f_0x^n + \cdots + f_ny^n \in K[x, y]$ a binary form of nonzero discriminant. Let $S_f \subset \mathbb{P}^1_K$ be the zero locus of f, $L_f = H^0(S_f, \mathcal{O}_{S_f})$, an étale K-algebra of rank n, and define $I_f = H^0(S_f, \mathcal{O}_{S_f}(n-3))$, $J_f = H^0(S_f, \mathcal{O}_{S_f}(n-2))$, as in §**??**. Then I_f, J_f are free L_f-modules of rank 1, although they do not have distinguished generators. We write $I_f^\times \subset I_f$ for the set of elements which generate I_f as an L_f-module; then $I_f^\times$ receives an action of $L_f^\times$. We define $J_f^\times$ similarly.

Definition 3.1. *We call an f-module a tuple (M, φ, e), where M is a free L_f-module of rank 1, φ : $M \otimes_{L_f} M \to I_f$ is an isomorphism of L_f-modules, and $e \in \wedge^n_K M$ is a non-zero element, satisfying the following property: let $\mathcal{B} = \{b_1, \dots, b_n\}$ be any K-basis of L_f such that $b_1 \wedge \cdots \wedge b_n = e$, and let A_x, A_y be the matrices of the symmetric K-bilinear forms $\mathrm{ev}_x \circ \varphi$, $\mathrm{ev}_y \circ \varphi$ on M. Then $\mathrm{disc}(xA_y - yA_x) = f(x, y)$.*

Henceforth we will write disc_e to denote discriminant with respect to some basis $\mathcal{B} = \{b_1, \dots, b_n\}$ such that $b_1 \wedge \cdots \wedge b_n = e$. Thus we can express the condition defining an f-module in the form $\mathrm{disc}_e(x\mathrm{ev}_y \circ \varphi - y\mathrm{ev}_x \circ \varphi) = f(x, y)$.

An isomorphism $(M, \varphi, e) \to (M', \varphi', e')$ of f-modules is an L_f-linear isomorphism $M \to M'$ intertwining the other data. We call (A_y, A_x) the pair associated to the f-module M and basis $\mathcal{B}$.

Proposition 3.2. *The map $(M, \varphi, e) \mapsto (A_y, A_x)$ determines a bijection between the following two sets:*

1. *The set of isomorphism classes of f-modules.*
2. *The set of orbits $\mathrm{SL}_n(K) \backslash V_f(K)$.*

Proof. It is clear that the map (1) → (2) is well-defined. To show that it is bijective, we construct an inverse. Consider a pair $(A, B) \in V_f(K)$. We construct a structure (M, φ, e) of f-module on $M = K^n$, together with basis $\mathcal{B} = \{b_1, \dots, b_n\}$ with $b_1 \wedge \cdots \wedge b_n = e$, such that (A, B) is the pair associated to (M, φ, e) and choice of basis $\mathcal{B}$.

Let $(s, -r) \in K^2$ be such that $sf(s, -r) \neq 0$, and let $a(x, y) = rx + sy$. We set $A_a = sA + rB$. Then $\det A_a \neq 0$, and we give M the unique L_f-module structure for which the element $x/a \in L_f$ (more precisely, the image of $x/a \in H^0(D^+(a), \mathcal{O}_{\mathbb{P}^1_K})$ in $H^0(S_f, \mathcal{O}_{S_f}) = L_f$) acts via the matrix $A_a^{-1}B$. We can then check the following:

- The element y/a acts on M via the matrix $A_a^{-1}A$. (Use the relation $r(x/a) + s(y/a) = 1$ in L_f.)
- The induced L_f-module structure on M is independent of the choice of a. (If b is another choice of linear form, use the relation $x/b = (b/a)^{-1}(x/a)$ in L_f and check that the two possible actions of x/b agree.)

We next define the pairing φ. Equivalently, we must define an isomorphism $M \cong \mathrm{Hom}_{L_f}(M, I_f)$ of L_f-modules. By duality (Proposition **??**), there is a canonical isomorphism $\mathrm{Hom}_{L_f}(M, J_f) \cong \mathrm{Hom}_K(M, K)$, given by composition with ζ. Multiplication by a defines an isomorphism $\times a : I_f \to J_f$. The pairing given by the symmetric matrix A_a gives an L_f-linear map $\psi_a : M \to \mathrm{Hom}_K(M, K) \cong \mathrm{Hom}_{L_f}(M, J_f)$. We define $\varphi = (\times a)^{-1} \circ \psi_a$. We can then check the following:

- The map φ is an isomorphism of L_f-modules, which is independent of the choice of a. (If b is another choice of linear form, we must check that $(b/a)\psi_a = \psi_b$. This is equivalent to the assertion that $A_a(b/a) = A_b$ as K-linear forms $M \times M \to K$, which follows from the formula for the action of b/a on $M = K^n$.)
- The matrices of $A_y = \mathrm{ev}_y \circ \varphi$ and $A_x = \mathrm{ev}_x \circ \varphi$ with respect to the standard basis $\mathcal{B} = \{b_1, \dots, b_n\}$ of K^n are equal to A and B, respectively. (This is essentially the same as the previous point, replacing b by y and using Lemma **??**.)

To complete our construction, we can therefore take $e = b_1 \wedge \cdots \wedge b_n$. It is clear that this gives a well-defined map (2) → (1) (i.e. well-defined at the level of equivalence classes), which is inverse to the map (1) → (2). This completes the proof. □

If (M, φ, e) is an f-module, and $u \in M$ is an L_f-module generator, then $\varphi(u \otimes u) \in I_f^\times$. Multiplication by u^{-1} determines an isomorphism $M \to L_f$ of L_f-modules, so we get an element $u_*^{-1}(e) \in \wedge_K^n L_f$. This gives a well-defined map from the set of isomorphism classes of f-modules to the set of equivalence classes in $I_f^\times \times (\wedge_K^n L_f - \{0\})$, where we define pairs (α, z) and (α', z') to be equivalent if there exists $\beta \in L_f^\times$ such that $(\alpha', z') = (\beta^2 \alpha, \mathbf{N}_{L_f/K}(\beta)^{-1} z)$.

Proposition 3.3. *The map just defined has the following properties:*

1. *It is injective.*
2. *Let $(\alpha, z) \in I_f^\times \times (\wedge_K^n L_f - \{0\})$. Then the equivalence class of (α, z) lies in the image if and only if $\operatorname{disc}_z(x(\operatorname{ev}_y \circ \varphi_\alpha) - y(\operatorname{ev}_x \circ \varphi_\alpha)) = f(x, y)$, where $\varphi_\alpha : L_f \otimes_{L_f} L_f \to I_f$ is defined by $\varphi_\alpha(a \otimes b) = ab\alpha$.*

Proof. Suppose that f-modules (M, φ, e) and (M', φ', e') have the same image. It follows that we can choose basis elements u, u' for M, M' respectively such that $\varphi(u \otimes u) = \varphi'(u' \otimes u')$ and $u_*^{-1}(e) = (u')_*^{-1}(e')$. Then the isomorphism $M \to M'$ of L_f-modules sending u to u' intertwines the other structures, so is an isomorphism of f-modules.

The characterisation of the image is essentially the definition of an f-module. □

We remark that if M is a free L_f-module of rank 1, $\varphi : M \otimes_{L_f} M \to I_f$ is an isomorphism, and $e \in \wedge_K^n M$ is non-zero, then $\operatorname{disc}_e(x(\operatorname{ev}_y \circ \varphi) - y(\operatorname{ev}_x \circ \varphi))$ is a $K^\times$-multiple of $f(x, y)$. To check equality, it is therefore enough to check that

$$\operatorname{disc}_e(s(\operatorname{ev}_y \circ \varphi) + r(\operatorname{ev}_x \circ \varphi)) = \operatorname{disc}_e(\operatorname{ev}_{rx+sy} \circ \varphi) = f(s, -r)$$

for some $(r, s) \in K^2$ such that $f(s, -r) \neq 0$.

Corollary 3.4. *Suppose that K is algebraically closed. Then the set $\mathrm{SL}_n(K) \backslash V_f(K)$ has exactly one element.*

Proof. Since L_f is an étale K-algebra and K algebraically closed, $L_f \simeq K \times \cdots \times K$, every element of L_f is a square and any two elements of $I_f^\times$ differ by a square in $L_f^\times$. Moreover, for every $\alpha \in I_f^\times$ there exists $e \in (\wedge_K^n L_f - \{0\})$ such that (L_f, φ_α, e) is an f-module, in the notation of Proposition **??**. If $e' \in (\wedge_K^n L_f - \{0\})$ is such that $(L_f, \varphi_\alpha, e')$ is an f-module, then $e' = \pm e$. Therefore every pair $(\alpha', z') \in I_f^\times \times (\wedge_K^n L_f - \{0\})$ is equivalent to the fixed pair (α, e). □

A common special case is when $f_0 \neq 0$ (so in particular $f(1, 0) \neq 0$). Then various objects become more explicit, because S_f is contained in the distinguished affine open $D^+(y) \subset \mathbb{P}_K^1$. Taking $X = x/y$, we find that $L_f = K[X]/(f(X, 1))$. Thus L_f has the power basis $1, X, \ldots, X^{n-1}$, while I_f, J_f have L_f-basis the elements y^{n-3}, y^{n-2}, respectively.

Lemma 3.5. *With notation as in the previous paragraph, the map $\zeta : J_f \to K$ is equal to $f_0^{-1}(x^{n-1}y^{-1})^*$, where $(y^{n-2})^*, (xy^{n-3})^*, \ldots, (x^{n-1}y^{-1})^*$ is the K-basis of $\operatorname{Hom}_K(J_f, K)$ dual to the K-basis y^{n-2}, $xy^{n-3}, \ldots, x^{n-1}y^{-1}$ of J_f.*

Proof. We compute using the Cech cohomology of the affine covering $U_0 = D^+(y)$, $U_1 = D^+(x)$ of $\mathbb{P}_K^1$. Our hypothesis $f_0 \neq 0$ is equivalent to the condition $S_f \subset D(y)$. Thus the map

$$H^0(S_f, \mathcal{O}_{S_f}(n-2)) \to H^0(U_0, \mathcal{O}_{S_f}(n-2)) = K[X] \cdot y^{n-2}/K[X] \cdot y^{-2} f(x, y) = K[X] \cdot y^{n-2}/f(X, 1)K[X] \cdot y^{n-2}$$

is an isomorphism. This shows that the elements $y^{n-2}, xy^{n-3}, \ldots, x^{n-1}y^{-1} \in K[X] \cdot y^{n-2}$ project to a K-basis of J_f. Since $(xy)^{-1} f(x, y) \in H^0(U_0 \cap U_1, \mathcal{O}_{\mathbb{P}_K^1}(n-2))$ maps to 0 in $H^0(U_0 \cap U_1, \mathcal{O}_{S_f}(n-2))$, the Cech cocycle in

$$H^0(U_0, \mathcal{O}_{S_f}(n-2)) \oplus H^0(U_1, \mathcal{O}_{S_f}(n-2))$$

corresponding to $x^{n-1}y^{-1}$ is $(x^{n-1}y^{-1}, (x^{n-1}y^{-1} - f_0^{-1}(xy)^{-1}f(x,y)))$.

Recall that ζ is defined as the composite of the boundary map $J_f = H^0(S_f, \mathcal{O}_{S_f}(n-2)) \to H^1(\mathbb{P}^1_K, \mathcal{O}_{S_f}(-2))$ associated to the short exact sequence

$$0 \to \mathcal{O}_{\mathbb{P}^1_K}(-2) \overset{\times f}{\to} \mathcal{O}_{\mathbb{P}^1_K}(n-2) \to \mathcal{O}_{S_f}(n-2) \to 0$$

with the isomorphism $H^1(\mathbb{P}^1_K, \mathcal{O}_{\mathbb{P}^1_K}(-2)) \cong K$ afforded by Serre duality. In particular, it vanishes on the codimension 1 subspace $H^0(\mathbb{P}^1_K, \mathcal{O}_{\mathbb{P}^1_K}(n-2))$ of J_f. To prove the lemma, we need to show that ζ takes the value f_0^{-1} on $x^{n-1}y^{-1}$.

The image of $x^{n-1}y^{-1}$ under the boundary homomorphism is represented by the cocycle

$$(x^{n-1}y^{-1} - (x^{n-1}y^{-1} - f_0^{-1}(xy)^{-1}f(x,y)))/f(x,y) = f_0^{-1}(xy)^{-1} \in H^0(U_0 \cap U_1, \mathcal{O}_{\mathbb{P}^1_K}(-2)).$$

By definition, the map ζ takes $(xy)^{-1}$ to 1, so it follows that $\zeta(x^{n-1}y^{-1}) = f_0^{-1}$, as claimed. □

In this way, we recover the description of the set $\mathrm{SL}_n(K)\backslash V_f(K)$ given in the special case $f_0 \neq 0$ in [**?**, Theorem 7]:

Corollary 3.6. *Suppose that $f_0 \neq 0$. Then the following three sets are in bijection:*

1. *The set of orbits $\mathrm{SL}_n(K)\backslash V_f(K)$.*
2. *The set of isomorphism classes of f-modules (M, φ, e).*
3. *The set of equivalence classes of pairs $(\alpha, z) \in L_f^\times \times K^\times$ satisfying $z^2\mathbf{N}_{L_f/K}(\alpha) = f_0^{n+1}$, where we say (α, z), (α', z') are equivalent if there exists $\beta \in L_f^\times$ such that $\alpha' = \beta^2\alpha$ and $z' = \mathbf{N}_{L_f/K}(\beta)^{-1}z$.*

Proof. Let $e = 1 \wedge X \wedge \cdots \wedge X^{n-1} \in \wedge^n_K L_f$. We define a bijection $L_f^\times \times K^\times \to I_f^\times \times (\wedge^n_K L_f - \{0\})$ by the formula

$$(\alpha, z) \mapsto (\alpha y^{n-3}, ze).$$

This descends to a bijection between the corresponding equivalence classes. If $\alpha \in L_f^\times$, let $\varphi_\alpha : L_f \otimes_{L_f} L_f \to I_f$ be the map $u \otimes v \mapsto \alpha uvy^{n-3}$.

What we need to check is that the condition $\mathrm{disc}_{ze}(\mathrm{ev}_y \circ \varphi_\alpha) = f_0$ arising from Proposition **??** is equivalent to the condition $z^2\mathbf{N}_{L_f/K}(\alpha) = f_0^{n+1}$ appearing in the statement of Corollary **??**. This is a computation:

$$\mathrm{disc}_{ze}(\mathrm{ev}_y \circ \varphi_\alpha) = z^2 \,\mathrm{disc}_e(\zeta(\alpha uvy^{n-2})) = z^2\mathbf{N}_{L_f/K}(\alpha)\,\mathrm{disc}_e(\zeta(uvy^{n-2})),$$

where we write e.g. $\zeta(uvy^{n-2})$ for the K-bilinear form $L_f \times L_f \to K$, $(u,v) \mapsto \zeta(uvy^{n-2})$. Using Lemma **??**, we see that $\zeta(uvy^{n-2}) = f_0^{-1}\tau(uv)$, where $\tau : L_f \to K$ is the element of the dual basis of the K-basis $1, X, \ldots, X^{n-1}$ of L_f corresponding to X^{n-1}. We have $\mathrm{disc}_e(\tau(uv)) = 1$, hence $\mathrm{disc}\,\zeta(uvy^{n-2}) = f_0^{-n}$, and finally

$$\mathrm{disc}_{ze}(\mathrm{ev}_y \circ \varphi_\alpha) = z^2\mathbf{N}_{L_f/K}(\alpha)f_0^{-n}.$$

This is what we needed. □

3.3 Integral orbits

We will use the point of view developed in §**??** to construct elements of $\mathrm{SL}_n(K)\backslash V_f(K)$ from K-rational divisors on hyperelliptic curves in §**??** below. We conclude with a lemma that will be useful in extending this construction to produce integral orbits. Since we now want to consider integral structures, let us reset our notation. Let A be a PID with fraction field K of characteristic 0, and let $f(x,y) = f_0x^n + \cdots + f_ny^n \in A[x,y]$ be a homogeneous polynomial of degree n and non-zero discriminant $\Delta(f) \in A$. Let $S_f \subset \mathbb{P}^1_A$ denote the zero locus of f, and let $S_{f,K}$ denote its generic fibre. Similarly, let $I_f = H^0(S_f, \mathcal{O}_{S_f}(n-3))$, $I_{f,K} = I_f \otimes_A K$ its generic fibre.

Lemma 3.7. *Let (M_K, φ_K, e) be an f-module, and suppose given an A-lattice $M \leq M_K$ satisfying the following conditions:*

1. *The restriction φ of the map $\varphi_K : M_K \times M_K \to I_{f,K}$ to $M \times M$ takes values in I_f.*
2. *There exists an A-basis $\mathcal{B}$ of M such that $\mathrm{disc}_{\mathcal{B}}(x(\mathrm{ev}_y \circ \varphi) - y(\mathrm{ev}_x \circ \varphi)) = uf(x,y)$ for some $u \in A^\times$.*

Then:

1. *There exists an A-basis $\mathcal{B}' = \{b'_1, \dots, b'_n\}$ of M such that $e = b'_1 \wedge \cdots \wedge b'_n$.*
2. *If A_y, A_x are the matrices of $\mathrm{ev}_y \circ \varphi$, $\mathrm{ev}_x \circ \varphi$ with respect to the basis $\mathcal{B}'$, then (A_y, A_x) is a representative in $V_f(A)$ for the $\mathrm{SL}_n(K)$-orbit corresponding to the f-module (M_K, φ_K, e) under Proposition* **??***.*

Proof. Let $\mathcal{C} = \{c_1, \dots, c_n\}$ be a K-basis of M_K such that $c_1 \wedge \cdots \wedge c_n = e$, so that

$$\mathrm{disc}_{\mathcal{C}}(x(\mathrm{ev}_y \circ \varphi) - y(\mathrm{ev}_x \circ \varphi)) = f(x,y).$$

If $g \in \mathrm{GL}_n(K)$ is the change of basis matrix for $\mathcal{B}$, $\mathcal{C}$, then $\mathrm{disc}_{\mathcal{B}} = \det(g)^2 \,\mathrm{disc}_{\mathcal{C}}$, hence $u = \det(g)^2$, and so $u = v^2$ for some $v \in A^\times$. It follows that we can choose e.g. $b'_1 = v^{-1}b_1$, $b'_i = b_i$ if $i = 2, \dots, n$. □

3.4 Reduction theory

We define a reduction covariant for the representation $V(\mathbb{Z})$ of $\mathrm{SL}_n(\mathbb{Z})$, a key technical tool in our reduction theory. To this end, we start with a slightly more general set-up.

Let M be an n-dimensional $\mathbb{R}$-vector space equipped with two symmetric bilinear forms $(-,-)_1, (-,-)_2 \colon M \times M \to \mathbb{R}$. Suppose that there exists a basis of M such that the Gram matrices A, B of $(-,-)_1, (-,-)_2$ have the property that $\det(Ax - By) \in \mathbb{R}[x,y]$ has nonzero discriminant. We explain how to associate to this data a canonical positive definite inner product H on M.

Lemma 3.8. *There exists a $\mathbb{C}$-basis $b_1, \dots, b_n$ of $M \otimes_{\mathbb{R}} \mathbb{C}$ such that the Gram matrices of $(-,-)_1, (-,-)_2$ are both diagonal. Moreover, such a basis is unique up to reordering and rescaling.*

Proof. It suffices to prove these claims when $M = \mathbb{R}^n$, in which case we may view the pair $(-,-)_1, (-,-)_2$ as an element $(A, B) \in V(\mathbb{R})$ with invariant binary form $f(x,y) \in \mathbb{R}[x,y]$ of nonzero discriminant. By Proposition **??** and its proof, M is part of an f-module (M, φ, e) with $\mathrm{ev}_y \circ \varphi = A$ and $\mathrm{ev}_x \circ \varphi = B$. Let $v \in M$ be an L_f-module generator. Let $e_1, \dots, e_n$ be the idempotents corresponding to a decomposition $L_f \otimes_{\mathbb{R}} \mathbb{C} \simeq \mathbb{C} \times \cdots \times \mathbb{C}$. Then $e_1 \cdot v, \dots, e_n \cdot v$ is a $\mathbb{C}$-basis of $M \otimes_{\mathbb{R}} \mathbb{C}$ and $\varphi(e_i \cdot v, e_j \cdot v) = e_ie_j\varphi(v,v) = 0$

if $i \neq j$, so the forms A, B are both diagonal in this basis, proving existence. To prove uniqueness (up to reordering and rescaling), let $(s, -r) \in \mathbb{R}^2$ with $sf(s, -r) \neq 0$, let $a(x, y) = rx + sy$ and let $A_a = sA + rB$. Then $\det(A_a) \neq 0$, and any basis in which A, B are diagonal has the property that $A_a^{-1}B$ is also diagonal. Since $\operatorname{disc}(f) \neq 0$, $A_a^{-1}B$ is regular semisimple, hence such a basis is unique up to reordering and rescaling. □

Construction 3.9. *Let $b_1, \ldots, b_n$ be a $\mathbb{C}$-basis of $M \otimes_{\mathbb{R}} \mathbb{C}$ for which the Gram matrices of $(-,-)_1, (-,-)_2$ are both diagonal. The assignment*

$$(b_i, b_j)_H = \begin{cases} \max(|(b_i, b_i)_1, |(b_i, b_i)|_2) & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

extends to a Hermitian form on $M \otimes_{\mathbb{R}} \mathbb{C}$. Let H be the restriction of this Hermitian form to M.

Proposition 3.10. *In the above construction, H is independent of the choice of $b_1, \ldots, b_n$, real valued and positive definite. Moreover, if $(M', (-,-)_1', (-,-)_2')$ is another such triple and $\psi\colon (M, (-,-)_1, (-,-)_2) \to (M', (-,-)_1', (-,-)_2')$ is an isomorphism of vector spaces intertwining the respective forms, then ψ intertwines H and H'.*

Proof. By Lemma **??**, the basis $b_1, \ldots, b_n$ is unique up to reordering and rescaling, and these operations do not affect H. Choosing the basis so that $b_1, \ldots, b_k \in M$ and $b_{k+1}, \ldots, b_n$ come in complex conjugate pairs shows that H is real valued. To show that H is positive definite, we must show that $\max(|(b_i, b_i)_1|, |(b_i, b_i)_2|) > 0$ for each $i = 1, \ldots, n$. This is true, since the nonzero discriminant condition implies ([**?**, Lemma 6.4(b)]) that either $(b_i, b_i)_1$ or $(b_i, b_i)_2$ is nonzero for each i. The claim about compatibility with isomorphisms follows from the fact that the definition of H is independent of any additional choices. □

Applying this construction to our framed vector space $W_{\mathbb{R}} = \mathbb{R}^n$ fixed in §**??**, we obtain an inner product $H_{A,B}$ on $W_{\mathbb{R}}$ associated to any pair $(A, B) \in V(\mathbb{R})$ of nonzero discriminant. Let X denote the set of equivalence classes of inner products on $W_{\mathbb{R}}$, where two inner products H, H' are defined to be equivalent when one is an $\mathbb{R}_{>0}$-multiple of the other. Write $[H] \in X$ for the image of an inner product H on $W_{\mathbb{R}}$. The assignment $(A, B) \to [H_{A,B}]$ defines a map

$$\mathcal{R}\colon V(\mathbb{R})^{\Delta \neq 0} \to X, \tag{3.1}$$

which we call the reduction covariant. The group $\mathrm{SL}_n(\mathbb{R})$ acts on X via the formula $(v, w)_{g \cdot H} = (g^{-1}v, g^{-1}w)_H$ or in terms of matrices, via the formula $g \cdot H = g^{-t}Hg^{-1}$. Let H_0 be the standard inner product on $W_{\mathbb{R}}$ with Gram matrix the identity matrix. The map $g \mapsto g \cdot H_0$ induces an $\mathrm{SL}_n(\mathbb{R})$-equivariant bijection $\mathrm{SL}_n(\mathbb{R})/\mathrm{SO}_n(\mathbb{R}) \simeq X$. The reduction covariant $\mathcal{R}$ is $\mathrm{SL}_n(\mathbb{R})$-equivariant by the last sentence of Proposition **??**, hence by passage to the quotient we obtain a map

$$\mathcal{R}\colon \mathrm{SL}_n(\mathbb{Z})\backslash V(\mathbb{Z})^{\Delta \neq 0} \to \mathrm{SL}_n(\mathbb{Z})\backslash X. \tag{3.2}$$

We record a computation of $\det H_{A,B}$ with respect to the standard basis of $W_{\mathbb{R}}$. Let $(A, B) \in V(\mathbb{R})$ with $\Delta(A, B) \neq 0$ and write

$$f_{A,B} = c \prod_{i=1}^{r} (x - \omega_i y) \prod_{j=1}^{k} (\eta_j x - y),$$

where $c, \omega_i, \eta_j \in \mathbb{C}$ and $|\omega_i|, |\eta_j| \leq 1$ for all i, j.

Lemma 3.11. *In the above notation,* $\det(H_{A,B}) = |c|$.

Proof. Let $c_0 \in \mathbb{C}$ be an element such that $c_0^n = (-1)^{n(n-1)/2}c$ and consider the pair of diagonal matrices $(A_0, B_0) \in V(\mathbb{C})$, where $A_0 = (c_0, \dots, c_0, c_0\eta_1, \dots, c_0\eta_k)$ and $B_0 = (c_0\omega_1, \dots, c_0\omega_r, c_0, \dots, c_0)$. A computation shows that $f_{A_0,B_0} = f_{A,B}$. By Corollary **??**, there exists a $g \in \mathrm{SL}_n(\mathbb{C})$ such that $g \cdot (A, B) = (A_0, B_0)$. By definition of the reduction covariant, $g^{-t}H_{A,B}(\bar{g})^{-1}$ is a diagonal matrix, with ith diagonal entry equal to the maximum of the absolute values of the ith diagonal entries of A_0 and B_0. Therefore $g^{-t}H_{A,B}(\bar{g})^{-1}$ is the diagonal matrix $(|c_0|, \dots, |c_0|)$, so $\det(H_{A,B}) = \det(g^{-t}H_{A,B}(\bar{g})^{-1}) = |c_0|^n = |c|$. □

We remark that the reduction covariant given above is not unique. For example, for either choice of $k = 1, 2$ we could carry out the analogue of Construction **??** with the formula

$$(b_i, b_j)_{H_k} = \begin{cases} |(b_i, b_i)_k| & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

to obtain a different $\mathrm{SL}_n(\mathbb{R})$-equivariant reduction covariant $\mathcal{R}_k : V(\mathbb{R})^{\Delta \neq 0} \to X$. Computation suggests that carrying out reduction theory using $\mathcal{R}_1$ will make the coefficients of A as small as possible (perhaps at the cost of making the coefficients of B large), and vice versa for $\mathcal{R}_2$, while performing reduction theory with $\mathcal{R}$ will tend to reduce the coefficients of both A and B. Our choice of $\mathcal{R}$ is made here so that we can prove Lemma **??**.

4 Construction of rational and integral orbits

We now connect the representation of SL_n on V to divisors on hyperelliptic curves. Let A be a PID of fraction field K of characteristic 0, let $g \geq 1$, and let $n = 2g + 2$. Let

$$f(x, y) = f_0x^n + f_1x^{n-1}y + \cdots + f_ny^n \in A[x, y]$$

be a homogeneous form of degree n such that the discriminant $\Delta(f) \in A$ is nonzero. Let $\pi : X \to \mathbb{P}^1_A$ denote the double cover described by the equation $z^2 = f(x, y)$, and $\iota : X \to X$ the associated involution. Then X_K is a smooth hyperelliptic curve. We write $j : S_f \hookrightarrow X$ for the closed subscheme defined by the equation $z = 0$. As suggested by the notation, S_f may be identified with the subscheme of $\mathbb{P}^1_A$ defined by the vanishing of $f \in H^0(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(n))$, studied in §**??**. We set $R_f = H^0(S_f, \mathcal{O}_{S_f})$ and $L_f = R_f \otimes_A K$.

Let $D_K \leq X_K$ be an effective divisor of degree $2g - 1$, not intersecting $S_{f,K}$. We can write $D_{\overline{K}} = \sum_{i=1}^{k}(\alpha_i : 1 : \beta_i) + \sum_{i=k+1}^{2g-1}(1 : 0 : \beta_i)$ for some numbers $\beta_i \in \overline{K}$. We set $w = \prod_{i=1}^{2g-1} \beta_i \in K$.

Let $U(x, y) = u_0x^{2g-1} + \cdots + u_{2g-1}y^{2g-1} \in A[x, y]$ be a primitive form of degree $2g - 1$ vanishing precisely at the points of $\pi(D_{\overline{K}})$ (with multiplicity). Then U is determined up to multiplication by $A^\times$. In this section, we will prove the following theorem.

Theorem 4.1. *Suppose that $f_0 \neq 0$, let $X = x/y$, and let $\alpha = U(X, 1)$ mod $f(X, 1) \in L_f = R_f \otimes_A K$. Let $w = \prod_{i=1}^{2g-1} \beta_i$. Then:*

1. *The pair $(\alpha, u_k^{-(g+1)} f_0^{n-1} w^{-1})$ corresponds, under the bijection of Corollary **??**, to an orbit in $\mathrm{SL}_n(K)\backslash V_f(K)$.*
2. *Furthermore, this orbit intersects $V_f(A)$ (i.e. has an integral representative).*

After the proof, we will give a more precise version of Theorem **??** (see Theorem **??**), which is what we will use in our applications. The rational orbit given by Theorem **??** could depend on the choice of U, although its orbit under the group $(\mathrm{SL}_n/\mu_2)(K)$ (through which the action of $\mathrm{SL}_n(K)$ factors) is insensitive to this choice.

We begin by proving the first part of Theorem **??**, which is a computation. Let $\omega_1, \dots, \omega_n \in \overline{K}$ be the roots of $f(X,1)$. Then $U(X,1) = u_k \prod_{i=1}^k (X - \alpha_i)$ and $f(X,1) = f_0 \prod_{j=1}^n (X - \omega_j)$. Since $f_0 \neq 0$, we have $w \in K^\times$, and $w^2 = f_0^{n-3-k} \prod_{i=1}^k \beta_i^2$.

We have

$$\begin{aligned}\mathbf{N}_{L_f/K}(\alpha) = \prod_{j=1}^n U(\omega_j, 1) = u_k^n \prod_{i=1}^k \prod_{j=1}^n (\omega_j - \alpha_i) \\ = u_k^n \prod_{i=1}^k \prod_{j=1}^n (\alpha_i - \omega_j) = u_k^n f_0^{-k} \prod_{i=1}^k f(\alpha_i, 1) = u_k^n f_0^{-k} \prod_{i=1}^k \beta_i^2 = u_k^n w^2 f_0^{3-n}.\end{aligned}$$

It follows that $(u_k^{-(g+1)} w^{-1} f_0^{n-1})^2 \mathbf{N}_{L_f/K}(\alpha) = f_0^{n+1}$. Looking at the statement of Corollary **??**, we see that we have indeed obtained an orbit in $\mathrm{SL}_n(K)\backslash V_f(K)$. In order to show that this orbit admits an integral representative, we first give a geometric construction for its corresponding f-module (in the sense of §**??**), in a way similar to [**?**, **?**]. Let $M_K = H^0(S_{f,K}, j_K^* \mathcal{O}_{X_K}(D_K))$. Then M_K is a free L_f-module of rank 1. There is an isomorphism of sheaves

$$\mathcal{O}_{X_K}(D_K) \otimes_{\mathcal{O}_{X_K}} \mathcal{O}_{X_K}(\iota^* D_K) \cong \mathcal{O}_{X_K}(2g-1) \tag{4.1}$$

given by $a \otimes b \mapsto abU$. After pullback along $j_K^* : S_{f,K} \hookrightarrow X_K$, this gives (using the identity $\iota \circ j = j$) an isomorphism of sheaves

$$j_K^* \mathcal{O}_{X_K}(D_K) \otimes_{\mathcal{O}_{S_{f,K}}} j_K^* \mathcal{O}_{X_K}(D_K) \cong j_K^* \mathcal{O}_{X_K}(2g-1),$$

hence an isomorphism

$$M_K \cong \mathrm{Hom}_{L_f}(M_K, I_f \otimes_A K)$$

of L_f-modules, which is determined up to $A^\times$-multiple by D_K (and determined completely by D_K and the choice of $U(x,y)$). Let $\varphi_K : M_K \otimes_{L_f} M_K \to I_f \otimes_A K$ denote the corresponding isomorphism. The first part of Theorem **??** says that (M_K, φ_K, e) is an f-module, where $e \in \wedge_K^n M_K$ is defined by the formula

$$e = u_0^{-(g+1)} f_0^g w^{-1} \cdot (\overline{1} \wedge X\overline{1} \wedge \cdots \wedge X^{n-1}\overline{1}),$$

and $\overline{1} \in M_K$ is the image of the tautological section $1 \in H^0(X_K, \mathcal{O}_{X_K}(D_K))$ under restriction. Indeed, our hypothesis that D_K and S_f do not intersect implies that $\overline{1}$ is an L_f-module generator for M_K.

We now proceed to construct an A-lattice $M \leq M_K$. We will rely heavily on the fact that X is an integral local complete intersection of dimension 2. Continuing to take D_K to be an effective divisor in X_K of degree $2g-1$, let us take D to be its Zariski closure in X, a closed subscheme of X which is A-flat of degree $2g-1$. Let $Z \subset X$ denote the intersection of the subscheme cut out by $U \in H^0(X, \mathcal{O}_X(2g-1))$ with the finitely many fibres of $X \to \operatorname{Spec} A$ above points where $\Delta(f)$ is not a unit. Then Z has codimension 2; let $V = X - Z$. Then $D_V = D \cap V$ is a Cartier divisor in V, and (by [**?**, Theorem 1.12]) the restriction functor $\mathrm{Refl}(X) \to \mathrm{Refl}(V)$ between the respective categories of reflexive coherent sheaves is an equivalence. Consequently, there is a unique (up to unique isomorphism) reflexive coherent sheaf, that we call $\mathcal{O}_X(D)$, on X, whose restriction to V is isomorphic to $\mathcal{O}_V(D_V)$. Moreover, since $D_V + \iota^* D_V = (U)_V$ as Cartier divisors, the isomorphism $\mathcal{O}_{X_K}(D_K) \cong \mathcal{H}om_{\mathcal{O}_{X_K}}(\mathcal{O}_{X_K}(\iota^* D), \mathcal{O}_{X_K}(2g-1))$ extends to an isomorphism

$$\mathcal{O}_X(D) \cong \mathcal{H}om_{\mathcal{O}_X}(\mathcal{O}_X(\iota^* D), \mathcal{O}_X(2g-1)) \tag{4.2}$$

of reflexive coherent sheaves of $\mathcal{O}_X$-modules (here we use that the sheaf Hom is itself reflexive, by [**?**, Corollary 1.18]).

Let $\mathcal{M} = j^* \mathcal{O}_X(D)$. We claim that $\mathcal{M}$ satisfies the hypotheses of Proposition **??**. Combining the result of this proposition with Lemma **??** will imply the truth of Theorem **??**. These hypotheses be checked locally on $\operatorname{Spec} A$, so we now assume that A is a DVR of uniformizer ϖ and residue field $k = A/(\varpi)$.

Lemma 4.2. *$\mathcal{M}_\eta$ is free of rank 1 at each generic point $\eta \in S_f$.*

Proof. By construction, $\mathcal{M}$ is locally free of rank 1 after restriction to $V = X - Z$. It suffices to check that the generic points of S_f all lie in V. Since V contains X_K, it certainly contains all generic points of S_f lying above the generic point of $\operatorname{Spec} A$. S_f has generic points lying above $\operatorname{Spec} k$ if and only if f is not primitive, in which case there is a generic point corresponding to the closed subscheme $\mathbb{P}^1_k$ of S_f. The intersection $\mathbb{P}^1_k \cap Z$ is finite (being contained in the zero set of $\overline{U}(x, y) \in k[x, y]$), so we see that we're done in this case too. □

Lemma 4.3. *The isomorphism of Equation (**??**) determines, by pullback, an isomorphism*

$$\mathcal{M} \cong \mathcal{H}om_{\mathcal{O}_{S_f}}(\mathcal{M}, \mathcal{O}_{S_f}(2g-1))$$

of sheaves of $\mathcal{O}_{S_f}$-modules.

Proof. We need to show that the natural map

$$j^* \mathcal{H}om_{\mathcal{O}_X}(\mathcal{O}_X(\iota^* D), \mathcal{O}_X(2g-1)) \to \mathcal{H}om_{\mathcal{O}_{S_f}}(j^* \mathcal{O}_X(\iota^* D), j^* \mathcal{O}_X(2g-1))$$

is an isomorphism. We can work on stalks. Let $p \in S_f$, and let u be a local equation for S_f at p. Applying the functor $\operatorname{Hom}_{\mathcal{O}_{X,p}}(\mathcal{O}_X(\iota^* D)_p, -)$ to the short exact sequence

$$0 \to \mathcal{O}_{X,p} \overset{\times u}{\to} \mathcal{O}_{X,p} \to \mathcal{O}_{S_f,p} \to 0,$$

we see that it is enough to know that $\operatorname{Ext}^1_{\mathcal{O}_{X,p}}(\mathcal{O}_X(\iota^* D)_p, \mathcal{O}_{X,p}) = 0$. This will follow from [**?**, Corollary 1.3] if we can check that $\mathcal{O}_X(\iota^* D)_p$ is Cohen–Macaulay. However, [**?**, Theorem 1.9] shows that since $\mathcal{O}_X(\iota^* D)_p$ is reflexive, it is S_2, hence Cohen–Macaulay. □

Lemma 4.4. *$H^1(S_f, \mathcal{M}) = 0$.*

Proof. If f is primitive, then S_f is affine, so there is nothing to prove. We therefore assume that $f(x, y) = \varpi^m g(x, y)$ for some $m \geq 1$, where $g(x, y) \in A[x, y]$ is primitive. Taking in hand the short exact sequence

$$0 \to \mathcal{O}_X(D)(-(g+1)) \overset{\times z}{\to} \mathcal{O}_X(D) \to \mathcal{M} \to 0,$$

we see that it is enough to show that $H^1(X, \mathcal{O}_X(D)) = 0$. By cohomology and base change, it is even enough to show that $H^1(X_k, \mathcal{L}) = 0$, where $\mathcal{L} = \mathcal{O}_X(D)|_{X_k}$. The curve X_k is the genus g 'hyperelliptic ribbon' [**?**, §1] given by the equation $z^2 = 0$, where $z \in \mathcal{O}_{\mathbb{P}^1_k}(-(g+1))$. Writing $\pi : X_k \to \mathbb{P}^1_k$ for the double cover, $\pi_* \mathcal{L}$ is a torsion free coherent sheaf, generically of rank 2, therefore a locally free sheaf of rank 2. (We justify the assertion that $\pi_* \mathcal{L}$ is torsion-free. By [**?**, Lemma 0AUV], it is equivalent to show that (0) is the unique associated prime of each stalk of $\pi_* \mathcal{L}$. By [**?**, Lemma 05DZ], it is equivalent to show that the unique associated prime of each stalk of $\mathcal{L}$ is the unique minimal prime. The sheaf $\mathcal{L}$ is S_1, since $\mathcal{O}_X(D)$ is S_2 and ϖ is a non-zero divisor in each stalk of $\mathcal{O}_X(D)$. The desired property therefore follows from [**?**, Lemma 031Q].) Since every vector bundle on $\mathbb{P}^1_k$ splits, we can find an isomorphism $\pi_* \mathcal{L} \cong \mathcal{O}(i) \oplus \mathcal{O}(j)$, where $i \leq j$ and (by considering the Euler characteristic) we have $i + j = g - 2$. We wish to show that $H^1(\mathbb{P}^1_k, \pi_* \mathcal{L}) = 0$, or equivalently that $i \geq -1$.

Multiplication by z gives a morphism $\pi_*\mathcal{L} \to \mathcal{O}_{\mathbb{P}^1_k}(g+1) \otimes \pi_*\mathcal{L}$, which can be represented as a matrix

$$z = \begin{pmatrix} R & S \\ T & Q \end{pmatrix} \in H^0 \begin{pmatrix} \mathcal{O}_{\mathbb{P}^1_k}(g+1) & \mathcal{O}_{\mathbb{P}^1_k}(i-j+g+1) \\ \mathcal{O}_{\mathbb{P}^1_k}(j-i+g+1) & \mathcal{O}_{\mathbb{P}^1_k}(g+1) \end{pmatrix}.$$

Since $i+j = g-2$, we have $i-j+g+1 = 2i+3$.

Suppose for contradiction that $i \leq -2$. Then $2i+3 < 0$, so $S = 0$. We can then calculate

$$z^2 = \begin{pmatrix} R^2 & 0 \\ (R+Q)T & Q^2 \end{pmatrix} = 0,$$

hence $R = Q = 0$. Since $H^0(X_k, \mathcal{L}) = H^0(\mathbb{P}^1_k, 0 \oplus \mathcal{O}(j))$, the matrix equality $z = \left(\begin{smallmatrix} 0 & 0 \\ T & 0 \end{smallmatrix} \right)$ shows that z annihilates $H^0(X_k, \mathcal{L})$.

Let $\overline{\eta}$ denote the generic point of X_k. Then $\mathcal{L}_{\overline{\eta}}$ is free of rank 1 over $\mathcal{O}_{X_k, \overline{\eta}}$ (since $\mathcal{O}_X(D)$ is locally free in a neighbourhood of $\overline{\eta}$), and generated by the image of the global section $1 \in H^0(X, \mathcal{O}_X(D))$. Since $\mathcal{L}_{\overline{\eta}}$ is not annihilated by z, this is a contradiction. □

We have now completed the proof of Theorem ??, and have in fact proved the following more precise statement:

Theorem 4.5. *With assumptions and notation as above, there exists an A-basis $\mathcal{B} = \{b_1, \dots, b_n\}$ for M with the following properties:*

1. *Let $e = b_1 \wedge \dots \wedge b_n$. Then (M_K, φ_K, e) is an f-module, corresponding under the bijection of Corollary ?? to the pair $(\alpha, u_k^{-(g+1)} f_0^{n-1} w^{-1}) \in L_f^\times \times K^\times$.*
2. *The matrices of the symmetric bilinear forms $\mathrm{ev}_y \circ \varphi_K$, $\mathrm{ev}_x \circ \varphi_K$ with respect to $\mathcal{B}$ have coefficients in A, and satisfy $\mathrm{disc}_e(x\mathrm{ev}_y - y\mathrm{ev}_x) = f(x,y)$.*
3. *Let $\overline{1} \in M$ denote the image of the section $1 \in H^0(X, \mathcal{O}_X(D))$ in $M = H^0(S_f, j^*\mathcal{O}_X(D))$. Then $\overline{1}$ is saturated, i.e. generates a saturated A-submodule of M.*

Proof. Only the last part remains to be proved. We can again localise and assume that A is a DVR of uniformizer ϖ, and must show that $\overline{1} \notin \varpi M$. We have constructed a pairing $M \times M \to I_f = H^0(S_f, \mathcal{O}_{S_f}(2g-1))$, and the pairing of $\overline{1}$ with itself is $U(x,y)$, which lies in the image of $H^0(\mathbb{P}^1_A, \mathcal{O}_{\mathbb{P}^1_A}(2g-1))$. This image is saturated, and $U(x,y)$ has unit content, by definition, so $U(x,y)$ is saturated in I_f. If $\overline{1}$ were a multiple of ϖ then $U(x,y)$ would be a multiple of ϖ^2, which is not the case. □

Let us now take $A = \mathbb{Z}$ and work out the consequences of all the theory developed so far. Fix a polynomial $f(x,y) = f_0x^n + \dots + f_ny^n \in \mathbb{Z}[x,y]$ of nonzero discriminant. Let X be the corresponding hyperelliptic curve over $\mathbb{Z}$, and let $D_{\mathbb{Q}}$ be an effective divisor on $X_{\mathbb{Q}}$ of degree $2g-1$ as at the beginning of §?? (therefore assumed not to intersect the locus $z = 0$). Let's assume further that $f_0 f_n \neq 0$. We choose a primitive polynomial $U(x,y) \in \mathbb{Z}[x,y]$ vanishing precisely at the points of $\pi(D_{\mathbb{Q}})$, with multiplicity; it is determined up to sign.

Let $M_{\mathbb{Q}} = H^0(S_{f,\mathbb{Q}}, j_{\mathbb{Q}}^*\mathcal{O}_{X_{\mathbb{Q}}}(D_{\mathbb{Q}}))$; then there is a natural isomorphism $\varphi_{\mathbb{Q}} : M_{\mathbb{Q}} \otimes_{L_f} M_{\mathbb{Q}} \to I_{f,\mathbb{Q}}$ of L_f-modules, satisfying $\varphi_{\mathbb{Q}}(\overline{1} \otimes \overline{1}) = U$, where $\overline{1} \in M_{\mathbb{Q}}$ is the restriction of the tautological section $1 \in H^0(X_{\mathbb{Q}}, \mathcal{O}_{X_{\mathbb{Q}}}(D_{\mathbb{Q}}))$.

The following result is a consequence of Theorem ??:

Proposition 4.6. *There exists a $\mathbb{Z}$-lattice $M \leq M_{\mathbb{Q}}$ with the following properties:*

1. There exists a $\mathbb{Z}$-basis of M with determinant e such that $(M_{\mathbb{Q}}, \varphi_{\mathbb{Q}}, e)$ is an f-module and $\varphi(M \otimes M) \leq I_f$.

2. $\bar{1} \in M$ is saturated.

The f-module structure on $M_{\mathbb{Q}}$ determines two symmetric bilinear forms $\mathrm{ev}_y \circ \varphi$ and $\mathrm{ev}_x \circ \varphi$ on $M_{\mathbb{Q}}$. Associated with these forms is the reduction covariant H on $M_{\mathbb{R}}$ from §**??**. We now compute the norm of the vector $\bar{1} \in M$ with respect to the reduction covariant. Let $X = x/y \in L_f$, and let $\omega_1, \ldots, \omega_n \in \mathbb{C}$ be the roots of $f(X, 1)$, which are nonzero by assumption. Order them so that $|\omega_1|, \ldots, |\omega_r| \leq 1$ and $|\omega_{r+1}|, \ldots, |\omega_n| > 1$. Write f_x for the partial derivative of f with respect to x, and define f_y similarly.

Proposition 4.7. *We have*

$$(\bar{1}, \bar{1})_H = \sum_{i=1}^{r} \frac{|U(\omega_i, 1)|}{|f_x(\omega_i, 1)|} + \sum_{i=r+1}^{n} \frac{|U(1, \omega_i^{-1})|}{|f_y(1, \omega_i^{-1})|}. \tag{4.3}$$

Proof. Let $e_i \in L_f \otimes \mathbb{C}$ be the idempotent corresponding to the root ω_i. Since $\bar{1}$ is an L_f-module generator, $e_1\bar{1}, \ldots, e_n\bar{1}$ is a $\mathbb{C}$-basis of $M_{\mathbb{C}}$. Moreover since the complexified form $\varphi\colon M_{\mathbb{C}} \times M_{\mathbb{C}} \to L_f \otimes \mathbb{C}$ is $L_f \otimes \mathbb{C}$-bilinear, $\varphi(e_i\bar{1}, e_j\bar{1}) = 0$ if $i \neq j$, hence the bilinear forms $\mathrm{ev}_x \circ \varphi$ and $\mathrm{ev}_y \circ \varphi$ have diagonal Gram matrices with respect to the basis $e_1\bar{1}, \ldots, e_n\bar{1}$. By Construction **??**, we have

$$(\bar{1}, \bar{1})_H = \sum_{i=1}^{n} \max(|\mathrm{ev}_y(\varphi(e_i\bar{1}, e_i\bar{1}))|, |\mathrm{ev}_x(\varphi(e_i\bar{1}, e_i\bar{1}))|).$$

We work out the summands explicitly. For $1 \leq i \leq n$ we have $\varphi(e_i\bar{1}, e_i\bar{1}) = e_i\varphi(\bar{1}) = e_iU$ by the L_f-bilinearity and the explicit description of φ from (**??**). We have $\mathrm{ev}_y(e_iU) = \zeta(ye_iU)$ by Lemma **??**. Using Lemma **??** and computing as in the proof of Corollary **??**, we have $\zeta(ye_iU) = f_0^{-1}\tau(e_iU(X, 1)) = U(\omega_i, 1)/f_x(\omega_i, 1)$. Similarly we have $\mathrm{ev}_x(\varphi(e_i\bar{1}, e_i\bar{1})) = \zeta(xe_iU) = \zeta((x/y)ye_iU) = \omega_i\zeta(ye_iU) = U(1, \omega_i^{-1})/f_y(1, \omega_i^{-1})$. It follows that

$$\max(|\zeta(ye_iU)|, |\zeta(xe_iU)|) = \begin{cases} |\zeta(ye_iU)| & |\omega_i| \leq 1 \\ |\zeta(xe_iU)| & |\omega_i| \geq 1. \end{cases}$$

If we order the roots so that $|\omega_1|, \ldots, |\omega_r| \leq 1$ and $|\omega_{r+1}|, \ldots, |\omega_n| > 1$, then we find

$$(\bar{1}, \bar{1})_H = \sum_{i=1}^{r} \frac{|U(\omega_i, 1)|}{|f_x(\omega_i, 1)|} + \sum_{i=r+1}^{n} \frac{|U(1, \omega_i^{-1})|}{|f_y(1, \omega_i^{-1})|}.$$

□

5 Equidistribution of the reduction covariant

The purpose of this section is to show that the reduction covariant (**??**) becomes equidistributed over integral orbits with respect to the natural measure on the space of lattices $\mathrm{SL}_n(\mathbb{Z})\backslash X$. This will follow from a modification of the geometry-of-numbers methods of [**?**, §4]. We closely follow the structure of [**?**, §3].

5.1 Preliminaries

Let $n = 2g + 2$ for some integer $g \geq 1$ and consider the action of $\mathrm{SL}(W) = \mathrm{SL}_n(\mathbb{Z})$ on the representation V from **??**.

Coordinates on $\mathrm{SL}_n(\mathbb{R})$. Let H_0 be the standard inner product on $W_{\mathbb{R}} = \mathbb{R}^n$, whose Gram matrix is the identity matrix. Then $K = \mathrm{SO}_{H_0}(\mathbb{R}) = \mathrm{SO}_n(\mathbb{R})$ is a maximal compact subgroup of $\mathrm{SL}_n(\mathbb{R})$. Let $T \leq \mathrm{SL}_n$ be the subgroup of diagonal matrices. For ease of notation, we will write $(t_1, \dots, t_n)$ for the element $\mathrm{diag}(t_1, \dots, t_n)$ of T. Let $N \leq \mathrm{SL}_n$ denote the subgroup of unipotent upper triangular matrices. By the Iwasawa decomposition, the product map $N(\mathbb{R}) \times T(\mathbb{R})^\circ \times K \to \mathrm{SL}_n(\mathbb{R})$ is a diffeomorphism.

Fundamental set for $\mathrm{SL}_n(\mathbb{Z})\backslash \mathrm{SL}_n(\mathbb{R})$. We will use the language of semialgebraic sets, see [**?**, Chapter 2]. We will use without further mention that semialgebraic sets are closed under finite unions, intersections, and (pre-)images under semialgebraic maps.

By definition, a Siegel set is a subset of $\mathrm{SL}_n(\mathbb{R})$ of the form $\omega \cdot T_c \cdot K$, where $\omega \subset N(\mathbb{R})$ is a relatively compact subset, $c \in \mathbb{R}_{>0}$ and $T_c \coloneqq \{(t_1, \dots, t_n) \in T(\mathbb{R})^\circ \colon t_1/t_2 > c, \dots, t_{n-1}/t_n > c\}$. (This definition of T_c differs from the one given in [**?**, §4.1.2] since our convention for the action of SL_n on V is slightly different, see §**??**.) For every such Siegel set, the set of $\gamma \in \mathrm{SL}_n(\mathbb{Z})$ with $\gamma \cdot \mathfrak{S} \cap \mathfrak{S} \neq \varnothing$ is finite [**?**, Corollaire 15.3] ('Siegel property'). Since SL_n is a Chevalley group, there exists a Siegel set $\mathfrak{S}$ with the property that $\mathrm{SL}_n(\mathbb{Z}) \cdot \mathfrak{S} = \mathrm{SL}_n(\mathbb{R})$. Explicitly, by [**?**, Chapter 4, Theorem 4.12], we can take any Siegel set with $c \leq 2/\sqrt{3}$ and ω containing all $n \in N(\mathbb{R})$ whose off-diagonal entries n_{ij} satisfy $|n_{ij}| \leq 1/2$. Fix such a $\mathfrak{S}$. After enlarging $\mathfrak{S}$, we may and do assume that ω, and consequently $\mathfrak{S}$, is open and semialgebraic.

The set $\mathfrak{S}$ will serve as our fundamental set for the left action of $\mathrm{SL}_n(\mathbb{Z})$ on $\mathrm{SL}_n(\mathbb{R})$. An $\mathrm{SL}_n(\mathbb{Z})$-orbit might be represented more than once in $\mathfrak{S}$, but this does not cause any problems as long as we incorporate the multiplicity function $m \colon \mathfrak{S} \to \mathbb{Z}_{\geq 1}$, defined by $m(x) = \#(\mathrm{SL}_n(\mathbb{Z}) \cdot x \cap \mathfrak{S})$. This function is bounded and has semialgebraic fibres.

Measures on $\mathrm{SL}_n(\mathbb{R})$ and X. Choose a generator of the rank-1 module of left invariant top differential forms of SL_n over $\mathbb{Z}$. This generator is unique up to sign and induces a bi-invariant Haar measure dg on $\mathrm{SL}_n(\mathbb{R})$. We equip the maximal compact subgroup K with its probability Haar measure. We now explain how these measures also induce measures on X and $\mathrm{SL}_n(\mathbb{Z})\backslash X$.

The standard inner product H_0 defines an element of X and the map $g \mapsto g \cdot H_0$ induces a $\mathrm{SL}_n(\mathbb{R})$-equivariant bijection $\mathrm{SL}_n(\mathbb{R})/K \simeq X$. We will use this identification without further mention. Since the measure dg on $\mathrm{SL}_n(\mathbb{R})$ is bi-invariant, it induces measures on $\mathrm{SL}_n(\mathbb{R})/K = X$ and $\mathrm{SL}_n(\mathbb{Z})\backslash \mathrm{SL}_n(\mathbb{R})/K = \mathrm{SL}_n(\mathbb{Z})\backslash X$. We denote the latter measure by μ. It is a standard fact that $\mu(\mathrm{SL}_n(\mathbb{Z})\backslash X)$ is finite; in fact, we have $\mu(\mathrm{SL}_n(\mathbb{Z})\backslash X) = \mathrm{vol}(\mathrm{SL}_n(\mathbb{Z})\backslash \mathrm{SL}_n(\mathbb{R})) = \zeta(2) \cdots \zeta(n)$.

Good subsets of X. Consider the surjective quotient map $\varphi \colon \mathfrak{S} \to \mathrm{SL}_n(\mathbb{Z})\backslash \mathrm{SL}_n(\mathbb{R})/K = \mathrm{SL}_n(\mathbb{Z})\backslash X$. We call a subset $U \subset \mathrm{SL}_n(\mathbb{Z})\backslash X$ *good* if it is relatively compact and $\varphi^{-1}(U) \subset \mathfrak{S}$ is a semialgebraic subset of $\mathrm{SL}_n(\mathbb{R})$. For example, the image under φ of a relatively compact semialgebraic subset of $\mathfrak{S}$ is good.

Lemma 5.1. *There exists a countable basis of good open subsets of* $\mathrm{SL}_n(\mathbb{Z})\backslash X$.

Proof. The proof is identical to that of [**?**, Lemma 3.2]. □

The next two lemmas ensure that good subsets have good properties for the purposes of the geometry-of-numbers arguments.

Lemma 5.2. *Let $U \subset \mathrm{SL}_n(\mathbb{Z})\backslash X$ be good and let $\bar{U} \subset \mathrm{SL}_n(\mathbb{R})$ be its preimage under the quotient map*

$\mathrm{SL}_n(\mathbb{R}) \to \mathrm{SL}_n(\mathbb{Z})\backslash X$. *Then for all* $h \in \mathrm{SL}_n(\mathbb{R})$ *we have*

$$\int_{g \in \mathfrak{S} \cap \bar{U} h} m(g)^{-1}\, dg = \mu(U).$$

Proof. Follows from pushing forward measures and the bi-invariance of dg. □

Lemma 5.3. *Let* $U \subset \mathrm{SL}_n(\mathbb{Z})\backslash X$ *be good, and let* $A \subset \mathrm{SL}_n(\mathbb{R})$ *be a compact semialgebraic subset. Let* $\bar{U}$ *be the preimage of* U *under the quotient map* $\mathrm{SL}_n(\mathbb{R}) \to \mathrm{SL}_n(\mathbb{Z})\backslash X$*. Then the set* $\{(g, h) \in \mathfrak{S} \times A \colon gh \in \bar{U}\}$ *is a semialgebraic subset of* $\mathrm{SL}_n(\mathbb{R}) \times \mathrm{SL}_n(\mathbb{R})$*.*

Proof. Identical to the proof of [**?**, Lemma 3.4]. □

Fundamental sets for $\mathrm{SL}_n(\mathbb{R})\backslash V(\mathbb{R})$**.** Let $B^s(\mathbb{R})$ denote the set of polynomials $f = f_0x^n + f_1x^{n-1}y + \cdots + f_ny^n \in \mathbb{R}[x]$ with nonzero discriminant. Let $I(m) \subset B^s(\mathbb{R})$ be the subset of polynomials having exactly $2m$ real roots. The sets $I(m)$ with $0 \le m \le n/2$ are the connected components of $B^s(\mathbb{R})$. Bhargava [**?**, §4.1.1] has partitioned the set of elements in $V(\mathbb{R})$ whose invariant form lies in $I(m)$ into components $V^{(m,\tau)}$ indexed by $1 \le \tau \le \lfloor 2^{2m-1} \rfloor$. Moreover, he constructs for each τ a fundamental set $L^{(m,\tau)}$ for the action of $\mathrm{SL}_n^{\pm}(\mathbb{R}) = \{g \in \mathrm{GL}_n(\mathbb{R}) \colon \det(g) = \pm 1\}$ on height-1 elements in $V^{(m,\tau)}$. This fundamental set is bounded and semialgebraic and has the property that the invariant binary form map $L^{(m,\tau)} \to \{f \in I(m) \colon \mathrm{Ht}(f) = 1\}$ is a semialgebraic isomorphism. We do not recall the precise definition of $L^{(m,\tau)}$ here, but we only mention that every element of $L^{(m,\tau)}$ is an $\mathbb{R}_{>0}$-multiple of an element (A, B), where A is the block matrix of the form

$$\big(\lambda_1, \cdots, \lambda_{2m}, \psi(\lambda_{2m+1}), \cdots, \psi(\lambda_{n/2+m})\big) \tag{5.1}$$

and B is the block matrix of the form

$$\big(\mu_1, \cdots, \mu_{2m}, \psi(\mu_{2m+1}), \cdots, \psi(\mu_{n/2+m})\big), \tag{5.2}$$

where $\psi(x + y\sqrt{-1}) = \begin{pmatrix} x & y \\ y & -x \end{pmatrix}$, $\lambda_i, \mu_i \in \mathbb{R}$ for all $1 \le i \le 2m$, $\lambda_i, \mu_i \in \mathbb{C}$ for all $2m + 1 \le i \le n/2 + m$ and $\max(|\lambda_i|, |\mu_i|) = 1$ for all $1 \le i \le n/2 + m$. (We note that there is a typo in [**?**, §4.1.1] in the definition of ψ since the displayed matrix there is not symmetric.)

Since we will be working with the group $\mathrm{SL}_n(\mathbb{R})$, we need to slightly modify the definition of $L^{(m,\tau)}$ when $m = 0$. Define $L^{(0,1+)} = L^{(0,1)}$ and $L^{(0,1-)} = g \cdot L^{(0,1)}$, where g is the diagonal matrix $(-1, 1, \cdots, 1)$. If $\tau = 1+$ (respectively $\tau = 1-$), let $V^{(0,\tau)}$ denote the set of those $(A, B) \in V(\mathbb{R})$ such that $f_{A,B} \in I(0)$ and the pair $(\alpha, z) \in L_f^{\times} \times \mathbb{R}^{\times}$ corresponding to the $\mathrm{SL}_n(\mathbb{R})$-orbit of (A, B) under Corollary **??** has the property that $z > 0$ (respectively $z < 0$). We have $V^{(0,1)} = V^{(0,1+)} \sqcup V^{(0,1-)}$. Define the indexing set

$$\Sigma = \{(0, 1+), (0, 1-)\} \cup \{(m, \tau) \colon 1 \le m \le n/2 \text{ and } 1 \le \tau \le 2^{2m-1}\}. \tag{5.3}$$

Lemma 5.4. *For each* $(m, \tau) \in \Sigma$, $L^{(m,\tau)}$ *is a fundamental set for the action of* $\mathrm{SL}_n(\mathbb{R})$ *on height-*1 *elements in* $V^{(m,\tau)}$*.*

Proof. Let $g \in \mathrm{SL}_n^{\pm}(\mathbb{R})$ be the diagonal matrix $(-1, 1, \cdots, 1)$. An orbit $\mathrm{SL}_n^{\pm}(\mathbb{R}) \cdot v \subset V(\mathbb{R})$ breaks up into one or two $\mathrm{SL}_n(\mathbb{R})$-orbits, depending on whether v is $\mathrm{SL}_n(\mathbb{R})$-conjugate to $g \cdot v$ or not. The explicit description of $L^{(m,\tau)}$ shows that if $m > 0$ then $g \cdot v = v$ for all $v \in L^{(m,\tau)}$, and so $L^{(m,\tau)}$ is a fundamental set for the $\mathrm{SL}_n(\mathbb{R})$-action on height 1 elements in $V^{(m,\tau)}$ since it was such a set for the $\mathrm{SL}_n^{\pm}(\mathbb{R})$-action. If $m = 0$ and $v \in L^{(0,1+)}$, then $g \cdot v \in L^{(0,1-)}$ is not $\mathrm{SL}_n(\mathbb{R})$-conjugate to v by Corollary **??**, so each $\mathrm{SL}_n^{\pm}(\mathbb{R})$-orbit in $V^{(0,1)}$ breaks up into two $\mathrm{SL}_n(\mathbb{R})$-orbits, corresponding to the decomposition $V^{(0,1)} = V^{(0,1+)} \sqcup V^{(0,1-)}$. □

Lemma 5.5. *If $(m,\tau) \in \Sigma$ and $v \in V^{(m,\tau)}$, then $\#\operatorname{Stab}_{\mathrm{SL}_n(\mathbb{R})}(v)$ equals $2^{n/2+m-1}$ if $m > 0$ and $2^{n/2}$ if $m = 0$.*

Proof. The index of $\operatorname{Stab}_{\mathrm{SL}_n(\mathbb{R})}(v)$ in $\operatorname{Stab}_{\mathrm{SL}_n^{\pm}(\mathbb{R})}(v)$ is 2 or 1, depending on whether $\mathrm{SL}_n^{\pm}(\mathbb{R}) \cdot v$ breaks up into one or two $\mathrm{SL}_n(\mathbb{R})$-orbits. This lemma therefore follows from (the proof of) Lemma **??** and the computation of $\#\operatorname{Stab}_{\mathrm{SL}_n^{\pm}(\mathbb{R})}(v)$ in [**?**, §3.2]. □

The next crucial lemma shows that the reduction covariant of §**??** is well behaved (in fact, constant) on the fundamental sets $L^{(m,\tau)}$.

Lemma 5.6. *For every pair (m,τ) as above and elements (A,B), $(A',B') \in L^{(m,\tau)}$, we have $[H_{A,B}] = [H_{(A',B')}]$.*

Proof. This follows from the explicit description of the elements in $L^{(m,\tau)}$ given by the expressions (**??**), (**??**). If $m = n/2$, then every pair $(A,B) \in L^{(m,\tau)}$ consists of diagonal matrices, and the condition $\max(|\lambda_i|, |\mu_i|) = 1$ above implies that $H_{A,B}$ equals (up to $\mathbb{R}_{>0}$-scaling) the standard inner product on $W_{\mathbb{R}} = \mathbb{R}^n$. The case that $m < n/2$ is similar; we omit the details. □

Since we are free to replace $L^{(m,\tau)}$ by $g \cdot L^{(m,\tau)}$ for some $g \in \mathrm{SL}_n(\mathbb{R})$ and preserve all of its required properties, we may and do assume that the $L^{(m,\tau)}$ have been chosen so that the reduction covariant of every element in $L^{(m,\tau)}$ equals the standard inner product $H_0 \in X$.

Counting lattice points. We recall the following proposition [**?**, Theorem 1.3], which strengthens a well-known result of Davenport [**?**].

Proposition 5.7. *Let $m, n \geq 1$ be integers, and let $Z \subset \mathbb{R}^{m+n}$ be a semialgebraic subset. For $T \in \mathbb{R}^m$, let $Z_T = \{x \in \mathbb{R}^n \colon (T,x) \in Z\}$, and suppose that all such subsets Z_T are bounded. Then*

$$\#(Z_T \cap \mathbb{Z}^n) = \operatorname{vol}(Z_T) + O(\max\{1, \operatorname{vol}(Z_{T,j})\}),$$

where $Z_{T,j}$ runs over all orthogonal projections of Z_T to all j-dimensional coordinate hyperplanes ($1 \leq j \leq n-1$). Moreover, the implied constant depends only on Z.

5.2 The equidistribution theorem

Define the height of a binary form $f(x,y) = f_0 x^n + \cdots + f_n y^n \in \mathbb{R}[x,y]$ by the formula $\operatorname{Ht}(f) = \max |f_i|$. Let $\mathcal{F}(X)$ be the set of integral binary forms $f = f_0 x^n + f_1 x^{n-1} y + \cdots + f_n y^n \in \mathbb{Z}[x,y]$ of nonzero discriminant and of height $< X$.

Lemma 5.8. *$\#\mathcal{F}(X) = 2^{n+1} X^{n+1} + O(X^n)$.*

Proof. We need to show that there are $O(X^n)$ polynomials with $\operatorname{Ht}(f) < X$ and $\Delta(f) = 0$. This follows immediately from [**?**, Lemma 3.1]. □

Define the height of an element $(A,B) \in V(\mathbb{R})$ to be the height of its invariant binary form: $\operatorname{Ht}(A,B) = \operatorname{Ht}(f_{A,B})$. We say an element $(A,B) \in V(\mathbb{Z})$ is degenerate if $\Delta(A,B) = \operatorname{disc}(f_{A,B}) = 0$, and nondegenerate otherwise. For any subset $S \subset V(\mathbb{Z})$, write S^{nd} for the subset of nondegenerate elements. Given $X \in \mathbb{R}_{>0}$

and an $\mathrm{SL}_n(\mathbb{Z})$-invariant subset $S \subset V(\mathbb{Z})$, let $N(S;X)$ be the number of nondegenerate $\mathrm{SL}_n(\mathbb{Z})$-orbits in S of height $< X$.

Bhargava [**?**, §4,Theorem 8] has determined asymptotics for the number of nondegenerate orbits in $V(\mathbb{Z})$ of bounded height under the action of $\mathrm{SL}_n^{\pm}(\mathbb{Z}) = \mathrm{GL}_n(\mathbb{Z})$. The same computation applies to nondegenerate $\mathrm{SL}_n(\mathbb{Z})$-orbits, showing that there exists a constant $c > 0$ such that

$$N(V(\mathbb{Z});X) = cX^{n+1} + o(X^{n+1}) \tag{5.4}$$

as $X \to +\infty$. The purpose of this section is to show, in analogy with [**?**, Theorem 3.8], that the reduction covariant (**??**) equidistributes over nondegenerate orbits, in the following sense:

Theorem 5.9. *Let $U \subset \mathrm{SL}_n(\mathbb{Z})\backslash X$ be a Borel-measurable subset whose boundary has measure zero. Then*

$$N(V(\mathbb{Z}) \cap \mathcal{R}^{-1}(U);X) = \frac{\mu(U)}{\mu(\mathrm{SL}_n(\mathbb{Z})\backslash X)} \cdot N(V(\mathbb{Z});X) + o(X^{n+1}).$$

The proof of Theorem **??** follows from a modification of the proof of Bhargava of the asymptotic (**??**) and is given below. We first prove a local version for good subsets. In the notation of §**??**, fix a pair $(m,\tau) \in \Sigma$ (see (**??**)), which corresponds to a subset $V^{(m,\tau)} \subset V(\mathbb{R})$, and write $V(\mathbb{Z})^{(m,\tau)} = V(\mathbb{Z}) \cap V^{(m,\tau)}$.

Let $\mathcal{J}$ be the rational nonzero constant of [**?**, Proposition 16]. Let $r_{m,n}$ equal $\frac{1}{2}\#\,\mathrm{Stab}_{\mathrm{SL}_n(\mathbb{R})}(v)$ for any $v \in V^{(m,\tau)}$; this constant is computed explicitly in Lemma **??**. For a subset S of $V(\mathbb{R})$ or $I(m)$, write $S_{<X}$ for the set of $s \in S$ with $\mathrm{Ht}(s) < X$. Equip $I(m)$ with the measure corresponding to the differential form $df_0 \wedge \cdots \wedge df_n$. Let

$$c_{m,\tau} = \mu(\mathrm{SL}_n(\mathbb{Z})\backslash X)\frac{|\mathcal{J}| \cdot \mathrm{vol}(I(m)_{<1})}{r_{m,n}},$$

where v denotes any element of $V^{(m,\tau)}$. Bhargava [**?**, Theorem 9] showed that

$$N(V(\mathbb{Z})^{(m,\tau)};X) = c_{m,\tau}X^{n+1} + o(X^{n+1}). \tag{5.5}$$

Proposition 5.10. *Let $U \subset \mathrm{SL}_n(\mathbb{Z})\backslash X$ be a good subset and let $S_U := V(\mathbb{Z})^{(m,\tau)} \cap \mathcal{R}^{-1}(U)$. Then*

$$N(S_U;X) = \frac{\mu(U)}{\mu(\mathrm{SL}_n(\mathbb{Z})\backslash X)}c_{m,\tau}X^{n+1} + o(X^{n+1}). \tag{5.6}$$

Proof. We will use the notations and choices made in §**??**. Write $\Lambda = \mathbb{R}_{>0}$, equipped with its multiplicative Haar measure $d^\times\lambda$, and let Λ act on $V(\mathbb{R})$ by scalar multiplication. Write $L = L^{(m,\tau)}$. Fix a compact, semialgebraic set $G_0 \subset \mathrm{SL}_n(\mathbb{R}) \times \Lambda$ of volume 1, with nonempty interior, that satisfies $K \cdot G_0 = G_0$ and whose projection onto Λ is contained in $[1, K_0]$ for some $K_0 > 1$.

For any subset $S \subset V(\mathbb{Z})^{(m,\tau)}$, averaging over G_0 [**?**, Equation (17)] shows

$$N(S;X) = \frac{1}{r_{m,n}}\int_{h\in G_0} \#[S \cap (\mathfrak{S}\Lambda hL)_{<X}]\,dh,$$

with the caveat that elements on the right hand side are weighted by a function similar to [**?**, §6.5, Equation (6.5)]. We use the above expression to define $N(S;X)$ for subsets $S \subset V(\mathbb{Z})^{(m,\tau)}$ that are not necessarily $\mathrm{SL}_n(\mathbb{Z})$-invariant. A change of variables trick [**?**, Equation (22)] shows that for every $S \subset V(\mathbb{Z})^{(m,\tau)}$ we then have

$$N(S;X) = \frac{1}{r_{m,n}}\int_{g\in\mathfrak{S}}\int_{\lambda\in\Lambda} \#[S \cap (g\lambda G_0 L)_{<X}]m(g)^{-1}\,dg\,d^\times\lambda,$$

with the caveat that an element $v \in S \cap (g\lambda G_0 L)$ on the right hand side is counted with multiplicity $\#\{h \in G_0 \colon v \in g\lambda h L\}$.

We let V^{cusp} be the subset of all $(A, B) \in V(\mathbb{R})$ such that the top left entries of A and B have absolute value < 1. By cutting off the cusp arguments [?, Proposition 13], $N(S_U; X) = N(S'_U; X) + o(X^{n+1})$, so it remains to estimate $N(S'_U; X)$.

We first make the reduction covariant more explicit. If $g \in \mathfrak{S}, h = (h', \lambda') \in G_0$ (with $h' \in \mathrm{SL}_n(\mathbb{R})$ and $\lambda' \in \Lambda$), $\lambda \in \Lambda$ and $\ell \in L$, the $\mathrm{SL}_n(\mathbb{R})$-equivariance and Λ-invariance of $\mathcal{R}$ imply that $\mathcal{R}(gh\lambda\ell) = gh'\mathcal{R}(\ell)$. We have chosen L so that $\mathcal{R}(\ell) = 1 \in \mathrm{SL}_n(\mathbb{R})/K$, so $gh'\mathcal{R}(\ell) = gh'$ in X. Write $\bar{U}$ for the preimage of U under the quotient map $\mathrm{SL}_n(\mathbb{R}) \to \mathrm{SL}_n(\mathbb{Z})\backslash X$. Then we conclude that $\mathcal{R}(gh\lambda\ell) \in U$ if and only if $gh' \in \bar{U}$.

Let $Z_1 = \{(g, h) \in \mathfrak{S} \times G_0 \mid gh' \in \bar{U}\}$, where we write h' for the projection of $h \in G_0$ onto $\mathrm{SL}_n(\mathbb{R})$. By Lemma **??**, Z_1 is semialgebraic. Let Z_2 be the graph of the action map $Z_1 \times \Lambda \times L \to V(\mathbb{R})$ sending (g, h, λ, ℓ) to $gh\lambda v$. Since this map is algebraic, Z_2 is semialgebraic. So is the projection Z_3 of Z_2 onto $\mathfrak{S} \times V(\mathbb{R})$. Let $Z_4 = \{(g, v, X) \in Z_3 \times \mathbb{R}_{>0} \mid \mathrm{Ht}(v) < X\}$; this is again semialgebraic. For every $g \in \mathfrak{S}$ and $X \in \mathbb{R}_{>0}$ the set $B(g, X) \coloneqq \{v \in V(\mathbb{R}) \mid (g, v, X) \in Z_4\}$ is semialgebraic and equals $(g\Lambda G_0 L)_{<X} \cap \mathcal{R}^{-1}(U)$. We view $B(g, X)$ as a multiset where $v \in B(g, X)$ has multiplicity $\#\{(\lambda, h) \in \Lambda \times G_0 \mid v \in g\lambda h L\}$. Then $B(g, X)$ is partitioned into finitely many semialgebraic subsets of constant multiplicity. Applying Proposition **??** to Z_4 shows that

$$\#[V(\mathbb{Z}) \cap (B(g, X) \setminus V^{\mathrm{cusp}})] = \mathrm{vol}(B(g, X) \setminus V^{\mathrm{cusp}}) + E(g, X),$$

where $E(g, X)$ is the error term. The proof now proceeds in an identical way to that of [?, Proposition 15]: estimating $E(g, X)$ and $\mathrm{vol}(B(g, X) \cap V^{\mathrm{cusp}})$ shows that

$$N(S_U; X) = \frac{1}{r_{m,n}} \int_{g \in \mathfrak{S}} \mathrm{vol}(B(g, X)) m(g)^{-1} \, dg + o(X^{n+1}). \tag{5.7}$$

The change of measure formula of [?, Proposition 16] shows that

$$\frac{1}{2^{m+n}} \int_{g \in \mathfrak{S}} \mathrm{vol}(B(g, X)) m(g)^{-1} \, dg = \frac{|\mathcal{J}|}{2^{m+n}} \mathrm{vol}(I(m)_{<X}) \int_{g \in \mathfrak{S}} \int_{h \in G_0} \mathbf{1}_{\{gh' \in \bar{U}\}} m(g)^{-1} \, dh \, dg, \tag{5.8}$$

where $\mathbf{1}_T$ denotes the indicator function of a set T, and where h' denotes the $\mathrm{SL}_n(\mathbb{R})$-component of h. By switching the order of integration, Lemma **??** and using $\mathrm{vol}(G_0) = 1$, we calculate that

$$\int_{g \in \mathfrak{S}} \int_{h \in G_0} \mathbf{1}_{\{gh' \in \bar{U}\}} m(g)^{-1} \, dh \, dg = \int_{h \in G_0} \int_{g \in \mathfrak{S}} \mathbf{1}_{\bar{U}h'^{-1}} m(g)^{-1} \, dg \, dh = \int_{h \in G_0} \mu(U) \, dh = \mu(U). \tag{5.9}$$

Combining (**??**), (**??**) and (**??**) shows that

$$N(S_U; X) = \mu(U) \frac{|\mathcal{J}|}{r_{m,n}} \mathrm{vol}(I(m)_{<X}) + o(X^{n+1}) = \mu(U) \frac{|\mathcal{J}|}{r_{m,n}} \mathrm{vol}(I(m)_{<1}) X^{n+1} + o(X^{n+1}),$$

as required. □

Proof of Theorem **??**. Let

$$\underline{\nu}(U) = \mu(\mathrm{SL}_n(\mathbb{Z})\backslash X) \liminf_{X \to \infty} \frac{N(V(\mathbb{Z}) \cap \mathcal{R}^{-1}(U); X)}{N(V(\mathbb{Z}); X)}$$

and $\bar{\nu}(U)$ be the same expression with lim inf replaced by lim sup. It suffices to prove that $\underline{\nu}(U) = \bar{\nu}(U) = \mu(U)$. If U is good, this follows from summing (**??**) over all $(m, \tau) \in \Sigma$, Equation (**??**), and the fact that $V^s(\mathbb{R})$ is partitioned into the subsets $V^{(m,\tau)}$. To prove the theorem for general U, we bootstrap from the case of good subsets (and the trivial case $U = \mathrm{SL}_n(\mathbb{Z})\backslash X$).

Let U° be the interior of U. Since U° is open, by Lemma **??** there exists an increasing sequence $(U_n)_{n\geq 1}$ of good subsets whose union is U°. For every $n \geq 1$ we have $\underline{\nu}(U_n) \leq \underline{\nu}(U^\circ)$. Since U_n is good, $\underline{\nu}(U_n) = \mu(U_n)$, hence $\mu(U_n) \leq \underline{\nu}(U^\circ)$ for all $n \geq 1$. By continuity of the measure, $\mu(U_n) \to \mu(U^\circ)$ as $n \to \infty$. We conclude that $\mu(U^\circ) \leq \underline{\nu}(U^\circ)$. Let $\bar{U}$ denote the closure of U. Since the complement of $\bar{U}$ equals the interior of the complement of U, the above argument also shows $\bar{\nu}(\bar{U}) \leq \mu(\bar{U})$.

In conclusion, we have shown that

$$\mu(U^\circ) \leq \underline{\nu}(U^\circ) \leq \underline{\nu}(U) \leq \bar{\nu}(U) \leq \bar{\nu}(\bar{U}) \leq \mu(\bar{U}).$$

By assumption, $\mu(U^\circ) = \mu(\bar{U}) = \mu(U)$, so all the inequalities are in fact equalities, and the theorem follows. □

5.3 A neighbourhood of the cusp

Let $\epsilon > 0$. Say an inner product H on $W_{\mathbb{R}} = \mathbb{R}^n$ has an ϵ-small vector if there exists a nonzero $v \in \mathbb{Z}^n$ such that

$$(v, v)_H^{1/2} < \epsilon \cdot (\det H)^{1/2n}.$$

This condition only depends on the image $\mathrm{SL}_n(\mathbb{Z}) \cdot [H]$ of H in $\mathrm{SL}_n(\mathbb{Z})\backslash X$, so it makes sense say that an element of $\mathrm{SL}_n(\mathbb{Z})\backslash X$ has an ϵ-small vector. We now show that such elements form a small measure subset of $\mathrm{SL}_n(\mathbb{Z})\backslash X$.

Recall from §**??** that $\mathfrak{S} = \omega T_c K$ and that $\varphi\colon \mathfrak{S} \to \mathrm{SL}_n(\mathbb{Z})\backslash X$ denotes the projection. After possibly enlarging $\mathfrak{S}$, we may and do assume that $c < 1$. Let $T(\epsilon) = \{(t_1, \ldots, t_n) \in T_c \colon t_1 > c^{n-1}/\epsilon\}$ and $U_\epsilon = \varphi(\omega T(\epsilon) K) \subset \mathrm{SL}_n(\mathbb{Z})\backslash X$.

Proposition 5.11. *1. The set U_ϵ is Borel-measurable, its boundary has measure zero, and $\mu(U_\epsilon) \to 0$ as $\epsilon \to 0$.*

2. If H is an inner product on $W_{\mathbb{R}}$ that admits an ϵ-small vector, then the image of H in $\mathrm{SL}_n(\mathbb{Z})\backslash X$ lies in U_ϵ.

Proof. Since the boundary of a semialgebraic set has strictly smaller dimension [**?**, Proposition 2.8.13] and $\varphi^{-1}(U_\epsilon)$ is semialgebraic, the boundary of $\varphi^{-1}(U_\epsilon)$ has measure zero, hence the same is true for U_ε. To show that $\mu(U_\epsilon) \to 0$ as $\epsilon \to 0$, it suffices to show that $\mathrm{vol}(\omega T(\epsilon) K) \to 0$ as $\epsilon \to 0$. This is true because $\mathrm{vol}(\mathfrak{S}) < \infty$ and $\cap_{\epsilon > 0}(\omega T(\epsilon) K) = 0$.

To prove the second part, we may assume that H has determinant 1 and so (since φ is surjective) that $H = g \cdot H_0 = g^{-t} H_0 g^{-1}$ for some $g = nt$ with $n \in \omega$ and $t = (t_1, \ldots, t_n) \in T_c$. Let $v \in W = \mathbb{Z}^n$ be a nonzero element with $(v, v)_H^{1/2} < \epsilon$. We will show that this implies $t_1 > c^{n-1}/\epsilon$. Write $e_1, \ldots, e_n$ for the standard basis of $\mathbb{Z}^n$ and let $f_i = g \cdot e_i$ for all i. Since $(v, w)_H = (g^{-1}v, g^{-1}w)_{H_0}$ for all $v, w \in W_{\mathbb{R}}$, the basis $f_1, \ldots, f_n$ is orthonormal with respect to $(-, -)_H$. Write $v = \sum_{i=1}^k m_i e_i$ where m_i are integers, $1 \leq k \leq m$ and $m_k \neq 0$. A computation reveals that $e_i \in t_i^{-1} f_i + \mathrm{span}_{\mathbb{R}}\{f_1, \ldots, f_{i-1}\}$ for each i, so $v \in m_k t_k^{-1} f_k + \mathrm{span}_{\mathbb{R}}\{f_1, \ldots, f_{k-1}\}$. By the orthonormality of the f_i and the fact that m_k is a nonzero integer,

$$(v, v)_H^{1/2} \geq |m_k| t_k^{-1} \geq t_k^{-1}.$$

Therefore $t_k > 1/\epsilon$. Since $(t_1, \ldots, t_n) \in T_c$, we have $t_i > c \cdot t_{i+1}$ for all $1 \leq i \leq k-1$, so $t_1 > c^{k-1} t_k > c^{k-1}/\epsilon > c^{n-1}/\epsilon$, as required. □

Combining Proposition **??** with the equidistribution theorem (Theorem **??**) and (**??**) immediately shows:

Corollary 5.12. *There exist a constant $c_1 > 0$ such that for every $\epsilon > 0$,*

$$\limsup_{X\to\infty} \frac{\#\{f \in \mathcal{F}(X) \colon \exists\, (A,B) \in V_f(\mathbb{Z}) \text{ and } w \in W \text{ with } (w,w)_{H_{A,B}}^{1/2} < \epsilon(\det H_{A,B})^{1/2n}\}}{\#\mathcal{F}(X)} \leq c_1 \cdot \mu(U_\epsilon). \tag{5.10}$$

Moreover, $\mu(U_\epsilon) \to 0$ as $\epsilon \to 0$.

6 A height lower bound for divisors

6.1 A density 1 family

Let $g \geq 1$. Recall from §**??** that for a real number $X > 0$, $\mathcal{F}(X)$ denotes the set of integral binary forms $f(x,y) \in \mathbb{Z}[x,y]$ of degree $n = 2g+2$, nonzero discriminant and height at most X, and that $\#\mathcal{F}(X) = (2X)^{n+1} + O(X^n)$.

For $\delta > 0$, let $\mathcal{F}_\delta(X)$ be the subset of $f = f_0x^n + \cdots + f_ny^n \in \mathcal{F}(X)$ satisfying the following properties:

1. $f(x,y)$ is an irreducible polynomial in $\mathbb{Q}[x,y]$.
2. We have $|\operatorname{disc}(f)| \geq X^{2n-2-\delta}$.

Note that the first condition implies that $f_0f_n \neq 0$.

These conditions cut out a density-1 family:

Proposition 6.1. *We have $\mathcal{F}_\delta(X) = (2X)^{n+1} + o(X^{n+1})$.*

Proof. We need to show that the number of $f \in \mathcal{F}(X)$ failing the first or second condition is $o(X^{n+1})$. For the first condition, this follows from Hilbert's irreducibility theorem; for the second, this is [**?**, Lemma 6.1]. □

The relevance of this family comes from the next elementary proposition. To state it, let $f \in \mathcal{F}(X)$ and write

$$f(x,y) = c\prod_{i=1}^{r}(x - \omega_i y)\prod_{j=1}^{k}(\eta_j x - y), \tag{6.1}$$

where $c, \omega_i, \eta_j \in \mathbb{C}$, $|\omega_i| \leq 1$ and $|\eta_j| < 1$ for all i, j. Write $f_x(x,y)$ for the partial derivative of $f(x,y)$ with respect to x, and similarly for $f_y(x,y)$.

Let $U(x,y) = u_0x^m + u_1x^{m-1}y + \cdots + u_my^m \in \mathbb{Z}[x,y]$ be an integral binary form of degree m. Write $h(U) = \log(\max(|u_0|, \ldots, |u_m|))$.

Proposition 6.2. *If $f \in \mathcal{F}_\delta(X)$ then, in the above notation, we have*

$$n \cdot \log\left(\sum_{i=1}^{r}\frac{|U(\omega_i,1)|}{|f_x(\omega_i,1)|} + \sum_{j=1}^{k}\frac{|U(1,\eta_j)|}{|f_y(1,\eta_j)|}\right) - \log|c| \leq n\cdot h(U) - (n+1)\log X + n\delta\log X + \kappa_{n,m}, \tag{6.2}$$

where $\kappa_{n,m}$ is a constant that only depends on n and m.

Proof. We have the formulae

$$\operatorname{disc}(f) = c^{2n-2} \prod_{1\le i<j\le r} (\omega_i - \omega_j)^2 \prod_{1\le i<j\le k} (\eta_i - \eta_j)^2 \prod_{1\le i\le r, 1\le j\le k} (\omega_i\eta_j - 1)^2,$$

$$f_x(\omega_i, 1) = c \prod_{j\neq i}(\omega_i - \omega_j) \prod_{j=1}^{k}(\omega_i\eta_j - 1) \text{ for each } i = 1, \dots, r.$$

Each term $|\omega_i - \omega_j|, |\eta_i - \eta_j|, |\omega_i\eta_j - 1|$ is ≤ 2. Therefore, $|\operatorname{disc}(f)|/|f_x(\omega_i, 1)|$ is $|c|^{2n-3}$ times a product of $n(n-1) - (n-1)$ terms that are each ≤ 2. Hence $|\operatorname{disc}(f)|/|f_x(\omega_i, 1)| \le |c|^{2n-3}2^{(n-1)^2}$. Since $|\operatorname{disc}(f)| \ge X^{2n-2-\delta}$ and $|U(\omega_i, 1)| \le (m + 1)\max(|u_0|, \cdots, |u_m|)$, we have

$$\frac{|U(\omega_i, 1)|}{|f_x(\omega_i, 1)|} \le (m + 1)\max(|u_i|)2^{(n-1)^2}|c|^{2n-3}/X^{2n-2-\delta}.$$

The same upper bound applies to $\frac{|U(1,\eta_j)|}{|f_y(1,\eta_j)|}$, and summing these bounds shows that the left hand side of (**??**) is at most

$$n \log(n(m + 1)2^{(n-1)^2}) + n \cdot h(U) + (n(2n - 3) - 1)\log|c| - n(2n - 2 - \delta)\log X. \qquad (6.3)$$

It remains to upper bound $\log|c|$, which we achieve by showing that the elements ω_i, η_j are not too far removed from the unit circle. Indeed, [**?**, Theorem 1] shows that

$$\prod_{i=1}^{r} |\omega_i|^{-1} \prod_{j=1}^{k} |\eta_j|^{-1} \le \frac{n}{|f_n/f_0|} \sum_{i=0}^{n} |f_i/f_0|^2.$$

On the other hand, $|c|^2 \prod|\omega_i| \prod|\eta_j| = |f_0||f_n|$ by (**??**), so

$$|c|^2 = |f_0||f_n| \prod_{i=1}^{r} |\omega_i|^{-1} \prod_{j=1}^{k} |\eta_j|^{-1} \le n \sum_{i=0}^{n} |f_i|^2 \le n(n + 1)X^2,$$

hence $\log|c| \le \frac{1}{2}\log(n(n+1)) + \log X$. Combining the latter upper bound with (**??**) concludes the proof. □

6.2 Height functions

Let $f \in \mathcal{F}(X)$ and let $\pi\colon X_f \to \mathbb{P}^1_{\mathbb{Q}}$ be the hyperelliptic curve described by the equation $z^2 = f(x, y)$. If D is an effective divisor on X_f of degree m, let $U(x, y) = u_0x^m + \cdots + u_my^m \in \mathbb{Z}[x, y]$ be a primitive binary form of degree m vanishing precisely at the points of $\pi(D)$ with multiplicity. Then U is uniquely determined up to sign and we may define

$$h(D) = h(U) = \log\max(|u_0|, \dots, |u_m|).$$

This is a 'naive height' on the set of effective divisors on X_f. For example, if D has degree 1, corresponding to a point $P \in X_f(\mathbb{Q})$, then $h(D)$ equals $h(\pi(P))$, where $h(\alpha)$ denotes the logarithmic normalized Weil height of an element $\alpha \in \mathbb{P}^1(\bar{\mathbb{Q}})$. More generally, by [**?**, Theorem VIII.5.9] we have:

Lemma 6.3. *If D is the Galois orbit of an algebraic point $P \in X_f(\bar{\mathbb{Q}})$ of degree m, then $|h(D) - mh(\pi(P))| \le m\log 2$.*

Theorem 6.4. *Suppose that $f \in \mathcal{F}_\delta(X)$ and that D is an effective divisor on X_f of degree $2g - 1$. Then there exists $(A, B) \in V_f(\mathbb{Z})$ and a primitive $w \in W = \mathbb{Z}^n$ such that*

$$n \cdot \log(w, w)_{H_{A,B}} - \log\det H_{A,B} \le n \cdot h(D) - (n + 1)\log X + n\delta\log X + \kappa_n,$$

where κ_n is a constant that only depends on n.

Proof. Since $f = f_0x^n + \cdots + f_ny^n \in \mathcal{F}_\delta(X)$ is irreducible, $f_0f_n \neq 0$ and D does not intersect the irreducible Weierstrass locus S_f. By Propositions **??**, **??** and Lemma **??**, there exists an (A, B) and a primitive $w \in W$ such that $n\log(w, w)_{H_{A,B}} - \det H_{A,B}$ equals the left hand side of (**??**), so we conclude using Proposition **??**. □

6.3 Proof of the main theorem

Let $\mathcal{F} = \cup_{X>0}\mathcal{F}(X)$ be the set of all binary forms $f(x, y) \in \mathbb{Z}[x, y]$ of degree $n = 2g + 2$ and nonzero discriminant, ordered by height. If $S \subset \mathcal{F}$ is a subset and $\alpha \in [0, 1]$, we say that S has density α if

$$\lim_{X\to\infty} \frac{\#(S \cap \mathcal{F}(X))}{\#\mathcal{F}(X)} = \alpha.$$

Theorem 6.5. *Let $\epsilon > 0$ be arbitrary, and let S_ϵ denote the set of $f \in \mathcal{F}$ such that every effective divisor D on X_f of odd degree $\leq 2g - 1$ satisfies*

$$h(D) \geq \left(1 + \frac{1}{2g + 2} - \epsilon\right) \log \mathrm{Ht} f. \tag{6.4}$$

Then S_ϵ has density 1.

Proof. Adding the degree-2 divisor $\pi^{-1}(\infty)$ to D has the effect of multiplying U by y^2, which does not affect the quantity $h(D)$. It therefore suffices to prove that the set $S'_\epsilon \subset S_\epsilon$ of $f \in \mathcal{F}$ satisfying (**??**) only for D of degree $2g - 1$ has density 1.

Let $\mathcal{F}^{\text{bad}}$ be the subset of those $f \in \mathcal{F}$ such that there exists a degree $2g - 1$ divisor D for which (**??**) does *not* hold; it suffices to prove that $\mathcal{F}^{\text{bad}} \subset \mathcal{F}$ has density zero. Fix $\delta \in (0, \epsilon)$. By Proposition **??**, it suffices to prove that $\#(\mathcal{F}^{\text{bad}} \cap \mathcal{F}_\delta(X)) = o(X^{2g+3})$.

Write $\epsilon_1 = \epsilon - \delta$. Let $\mathcal{F}(X)^{\text{short}}$ be the subset of $f \in \mathcal{F}(X)$ for which there exists elements $(A, B) \in V_f(\mathbb{Z})$ and $w \in W - \{0\}$ that satisfy

$$n \log(w, w)_{H_{A,B}} - \log \det H_{A,B} < -\epsilon_1 \log X. \tag{6.5}$$

Theorem **??** implies that $\mathcal{F}^{\text{bad}} \cap \mathcal{F}_\delta(X) \subset \mathcal{F}(X)^{\text{short}}$ for sufficiently large X. The equidistribution theorem, more specifically Corollary **??**, shows that $\#\mathcal{F}(X)^{\text{short}} = o(X^{2g+3})$. Therefore $\#(\mathcal{F}^{\text{bad}} \cap \mathcal{F}_\delta(X)) = o(X^{2g+3})$, as required. □

This theorem and Lemma **??** immediately imply:

Corollary 6.6. *Let $\epsilon > 0$ be arbitrary. Then for* 100% *of $f \in \mathcal{F}(X)$, every algebraic point $P \in X_f(\bar{\mathbb{Q}})$ of odd degree $m \leq 2g - 1$ satisfies*

$$m \cdot h(\pi(P)) \geq \left(1 + \frac{1}{2g + 2} - \epsilon\right) \log \mathrm{Ht}(f).$$

References

[1] F. Barroero and M. Widmer. Counting lattice points and O-minimal structures. *Int. Math. Res. Not. IMRN*, (18):4932–4957, 2014.

[2] D. Bayer and D. Eisenbud. Ribbons and their canonical embeddings. *Trans. Amer. Math. Soc.*, 347(3):719–756, 1995.

[3] M. Bhargava. The geometric sieve and the density of squarefree values of invariant polynomials. arXiv preprint, `1402.0031v1`, 2014.

[4] M. Bhargava. A positive proportion of plane cubics fail the Hasse principle. arXiv preprint, `1402.1131v1`, 2014.

[5] M. Bhargava. Most hyperelliptic curves over $\mathbb{Q}$ have no rational points. arXiv preprint, `1308.0395v1`, 2015.

[6] M. Bhargava and B. H. Gross. The average size of the 2-Selmer group of Jacobians of hyperelliptic curves having a rational Weierstrass point. In *Automorphic representations and L-functions*, volume 22 of *Tata Inst. Fundam. Res. Stud. Math.*, pages 23–91. Tata Inst. Fund. Res., Mumbai, 2013.

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

[8] M. Bhargava, B. H. 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(2):451–493, 2017. With an appendix by Tim Dokchitser and Vladimir Dokchitser.

[9] M. Bhargava, A. Shankar, and X. Wang. Squarefree values of polynomial discriminants I. *Invent. Math.*, 228(3):1037–1073, 2022.

[10] M. Bhargava, A. Shankar, and X. Wang. Squarefree values of polynomial discriminants II. arXiv preprint, `2207.05592v1`, 2022.

[11] J. Bochnak, M. Coste, and M.-F. Roy. *Real algebraic geometry*, volume 36 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1998. Translated from the 1987 French original, Revised by the authors.

[12] A. Borel. *Introduction aux groupes arithmétiques.* Publications de l'Institut de Mathématique de l'Université de Strasbourg, XV. Actualités Scientifiques et Industrielles, No. 1341. Hermann, Paris, 1969.

[13] B. Conrad. *Grothendieck duality and base change*, volume 1750 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2000.

[14] J. E. Cremona, T. A. Fisher, and M. Stoll. Minimisation and reduction of 2-, 3- and 4-coverings of elliptic curves. *Algebra Number Theory*, 4(6):763–820, 2010.

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

[16] R. Hartshorne. *Residues and duality.* Lecture Notes in Mathematics, No. 20. Springer-Verlag, Berlin-New York, 1966. Lecture notes of a seminar on the work of A. Grothendieck, given at Harvard 1963/64, With an appendix by P. Deligne.

[17] R. Hartshorne. Generalized divisors on Gorenstein schemes. In *Proceedings of Conference on Algebraic Geometry and Ring Theory in honor of Michael Artin, Part III (Antwerp, 1992)*, volume 8, pages 287–339, 1994.

[18] J. Laga. Arithmetic statistics of Prym surfaces. *Math. Ann.*, 386(1-2):247–327, 2023.

[19] J. Laga and J. Thorne. 100% of odd hyperelliptic Jacobians have no rational points of small height. arXiv preprint `2405.10224v1`, 2024.

[20] J. Nakagawa. Binary forms and orders of algebraic number fields. *Invent. Math.*, 97(2):219–235, 1989.

[21] V. Platonov and A. Rapinchuk. *Algebraic groups and number theory*, volume 139 of *Pure and Applied Mathematics*. Academic Press, Inc., Boston, MA, 1994. Translated from the 1991 Russian original by Rachel Rowen.

[22] B. Poonen and E. Rains. Random maximal isotropic subspaces and Selmer groups. *J. Amer. Math. Soc.*, 25(1):245–269, 2012.

[23] B. Poonen and M. Stoll. A local-global principle for densities. In *Topics in number theory (University Park, PA, 1997)*, volume 467 of *Math. Appl.*, pages 241–244. Kluwer Acad. Publ., Dordrecht, 1999.

[24] J. H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, Dordrecht, second edition, 2009.

[25] K. Soundararajan. Equidistribution of zeros of polynomials. *Amer. Math. Monthly*, 126(3):226–236, 2019.

[26] M. Stoll and J. E. Cremona. On the reduction theory of binary forms. *J. Reine Angew. Math.*, 565:79–99, 2003.

[27] A. A. Swaminathan. 2-Selmer groups, 2-class groups, and the arithmetic of binary forms. Available at `http://arks.princeton.edu/ark:/88435/dsp012v23vx55w`, 2022.

[28] The Stacks project authors. The Stacks project. `https://stacks.math.columbia.edu`, 2024.

[29] J. A. Thorne. A remark on the arithmetic invariant theory of hyperelliptic curves. *Math. Res. Lett.*, 21(6):1451–1464, 2014.

[30] J. A. Thorne. Reduction theory for stably graded Lie algebras, 2023. Preprint.

[31] M. M. Wood. Rings and ideals parameterized by binary n-ic forms. *J. Lond. Math. Soc. (2)*, 83(1):208–231, 2011.

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