JOURNAL
de Théorie des Nombres
de BORDEAUX
anciennement Séminaire de Théorie des Nombres de Bordeaux

Martí OLLER

The density of ADE families of curves having squarefree discriminant

Tome 37, nº 3 (2025), p. 989-1029.
https://doi.org/10.5802/jtnb.1349

Le Journal de Théorie des Nombres de Bordeaux *est membre du Centre Mersenne pour l'édition scientifique ouverte*
http://www.centre-mersenne.org/
e-ISSN : 2118-8572

Journal de Théorie des Nombres
de Bordeaux **37** (2025), 989–1029

The density of ADE families of curves having squarefree discriminant

par Martí OLLER

Résumé. Nous déterminons la densité des courbes à discriminant sans facteur carré dans certaines familles de courbes qui proviennent des représentations de Vinberg, en montrant que la densité globale est le produit des densités locales. Pour ce faire, nous travaillons dans le contexte des thèses de doctorat de Thorne et Laga et utilisons les techniques de comptage d'orbites de Bhargava. Cet article généralise un résultat précédent de Bhargava, Shankar et Wang.

Abstract. We determine the density of curves having squarefree discriminant in some families of curves that arise from Vinberg representations, showing that the global density is the product of the local densities. We do so using the framework of Thorne and Laga's PhD theses and Bhargava's orbit -counting techniques. This paper generalises a previous result by Bhargava, Shankar and Wang.

1. Introduction

The aim of this paper is to determine the density of curves in certain families that have squarefree discriminant. We do so following the techniques in arithmetic statistics developed by Bhargava and his collaborators. The main idea is that many arithmetic objects of interest can be parametrised by the rational or integral orbits of a certain representation (G, V): in this situation, Bhargava's geometry-of-numbers methods allow to count these integral orbits of V, which consequently provides information on the desired arithmetic objects that would be otherwise difficult to obtain. This idea has led to many impressive results in number theory; see [3] or [16] for an overview.

The present paper is inspired by the recent paper [7] by Bhargava, Shankar and Wang, in which they compute the density of monic integral polynomials of a given degree that have squarefree discriminant. The main

Manuscrit reçu le 20 août 2024, révisé le 10 mars 2025, accepté le 9 mai 2025.
2020 *Mathematics Subject Classification.* 11N35, 11G30, 17B70.
Mots-clefs. Arithmetic statistics, geometry-of-numbers, squarefree sieve, Vinberg representations, graded Lie algebras.
The project that gave rise to these results received the support of a fellowship from "la Caixa" Foundation (ID 100010434). The fellowship code is LCF/BQ/EU21/11890111. The author wishes to thank them, as well as the Cambridge Trust and the DPMMS, for their support.

technical difficulty is to bound the tail estimate of polynomials having discriminant "weakly divisible" by the square a large prime (this notion will be defined later). They do so using the representation of $G = \mathrm{SO}_n$ on the space V of $n \times n$ symmetric matrices. By relating polynomials with discriminant divisible by p^2 for a large p to certain integral orbits of the representation (G, V), they get the desired result using the aforementioned geometry-of-numbers techniques. Similar methods were used in [8] in the non-monic case with a different representation, and also in [5] for certain families of elliptic curves (in particular, their F_2 case essentially corresponds to our D_4 case).

A key observation, which motivates our results, is that the representation studied in [7] arises as a particular case of the more general families of representations studied in [37]. These representations arise from Vinberg theory, which is the study of the invariant theory of representations arising from graded Lie algebras. Given a $\mathbb{Z}/2\mathbb{Z}$-grading of a simple Lie algebra with a simply laced Dynkin diagram (that is, of type A_n, D_n or E_n), we obtain a coregular representation (G, V), which in the A_n case corresponds to the representation studied in [7]. On the other hand, simply laced Dynkin diagrams also parametrise simple surface singularities. Grothendieck conjectured that there should be a representation-theoretic construction of this correspondence, and the first detailed proofs of this result were given by Esnault [14] and Slodowy [34]. Thorne combined Vinberg and Slodowy's approach to construct a family of curves by looking at the transverse slice of the so-called subregular orbit in V. The central fibre of this transverse slice (namely, where all invariant polynomials vanish) corresponds to a simple surface singularity of type ADE, the same type as the Lie algebra we started with. The family of curves that Thorne obtains is a semiuniversal deformation of this central fibre.

Arithmetic information about the constructed families of curves can be read off the constructed representation (G, V). Most notably, these representations have been used (implicitly or explicitly) to study the average size of the 2-Selmer groups of the Jacobians of these curves: see [4, 20, 28, 30, 33, 38] for some particular cases. Later, Laga unified, reproved and extended all these results in [21] in a uniform way, using an impressive geometric construction with the compactified Jacobian.

Our aim is to compute the density of curves having squarefree discriminant in these families of ADE curves. We will do so by reinterpreting the methods in [7] in the language of [21] and [37]. As a corollary, we will obtain the asymptotics for the number of integral reducible orbits of these representations, following [31].

Let $\mathcal{D}$ be a Dynkin diagram of type A, D, E. In Section 2.1, we will construct a representation (G, V) associated to $\mathcal{D}$, and in Section 2.3 we will

construct a family of curves $C \to B$. Here, B is isomorphic to the Geometric Invariant Theory (GIT) quotient $V /\!\!/ G := \operatorname{Spec} \mathbb{Q}[V]^G$. We see that B can be identified with an affine space, and we write $B = \operatorname{Spec} \mathbb{Q}[p_{d_1}, \dots, p_{d_r}]$. Given $b \in B$, we define its *height* to be

$$\operatorname{ht}(b) := \sup \left(|p_{d_1}(b)|^{1/d_1}, \dots, |p_{d_r}(b)|^{1/d_r} \right).$$

Denote by C_b the preimage of a given $b \in B$ under the map $C \to B$; it will be a curve of the form given by Table 2.2. The main result of this paper concerns the density of squarefree values of the discriminant $\Delta(C_b)$ of the curve (or equivalently, the discriminant $\Delta(b)$ defined in Section 2.1). A definition for the discriminant of a plane curve can be found in [36, Section 2], for instance. We remark that in our definition of discriminant, we assume that it is an integer-valued polynomial in multiple variables, normalised so that the coefficients have greatest common divisor 1 (for instance, the usual discriminant for elliptic curves contains a factor of 16: we omit it in our case).

Our result is related to the p-adic density of these squarefree values: we will denote by $\rho(\mathcal{D}_p)$ the p-adic density of curves in the family $C \to B$ having discriminant indivisible by p^2 in $\mathbb{Z}_p$; this is obtained by taking all the (finitely many) elements in $b \in B(\mathbb{Z}/p^2\mathbb{Z})$ and counting the proportion of them that have non-zero discriminant in $\mathbb{Z}/p^2\mathbb{Z}$. We note that under our assumptions on the discriminant, none of the local densities vanish; this can be checked with a case-by-case computation.

Theorem 1.1. *We have*

$$\lim_{X \to \infty} \frac{\#\{b \in B(\mathbb{Z}) \mid \Delta(b) \text{ is squarefree, } \operatorname{ht}(b) < X\}}{\#\{b \in B(\mathbb{Z}) \mid \operatorname{ht}(b) < X\}} = \prod_p \rho(\mathcal{D}_p).$$

To prove this theorem, we need to obtain a tail estimate to show that not too many $b \in B(\mathbb{Z})$ have discriminant divisible by m^2 for large squarefree integers m. A key observation in [7] is to separate those $b \in B(\mathbb{Z})$ with $p^2 | \Delta(b)$ for a prime p in two separate cases:

(1) If $p^2 | \Delta(b + pc)$ for all $c \in B(\mathbb{Z})$, we say p^2 *strongly divides* $\Delta(b)$ (in other words, p^2 divides $\Delta(b)$ for "mod p reasons").
(2) If there exists $c \in B(\mathbb{Z})$ such that $p^2 \nmid \Delta(b + pc)$, we say p^2 *weakly divides* $\Delta(b)$ (in other words, p^2 divides $\Delta(b)$ for "mod p^2 reasons").

Similarly, given a squarefree number m, we say that m^2 strongly (resp. weakly) divides $\Delta(b)$ if p^2 strongly (resp. weakly) divides $\Delta(b)$ for all primes p dividing m. We let $\mathcal{W}_m^{(1)}, \mathcal{W}_m^{(2)}$ denote the set of $b \in B(\mathbb{Z})$ whose discriminant is strongly (resp. weakly) divisible by m^2. We prove tail estimates for these two sets separately. The argument in the weakly divisible case will require our squarefree integers m to avoid a finite number of primes: in Section 2.4, we will consider an integer N which contains all "bad primes".

Theorem 1.2. *There exists a constant $\delta > 0$ such that for any positive real number M we have:*

$$\sum_{\substack{m>M \\ m \text{ squarefree}}} \left\{ b \in \mathcal{W}_m^{(1)} \,\middle|\, \mathrm{ht}(b) < X \right\} = O_\varepsilon\left(\frac{X^{\dim V+\varepsilon}}{M} \right) + O_\varepsilon\left(X^{\dim V-1+\varepsilon} \right),$$

$$\sum_{\substack{m>M \\ m \text{ squarefree} \\ (m,N)=1}} \left\{ b \in \mathcal{W}_m^{(2)} \,\middle|\, \mathrm{ht}(b) < X \right\} = O_\varepsilon\left(\frac{X^{\dim V+\varepsilon}}{M} \right) + O\left(X^{\dim V-\delta} \right).$$

The implied constants are independent of X and M.

As in [7, Theorem 1.5 (a)], the strongly divisible case follows from the use of the Ekedahl sieve; more precisely, it follows from the results in [2, Theorem 3.5, Lemma 3.6] and the fact that the discriminant polynomial is irreducible by [21, Lemma 4.2]. Therefore, it remains to prove the (substantially harder) weakly divisible case, which is the content of most of this paper.

We start in Section 2 by giving the necessary background and introducing our objects of interest, most importantly the representation (G, V) coming from Vinberg theory and the associated family of curves $C \to B$. The main step in the proof of Theorem 1.2 is done in Section 3, where given a $b \in \mathcal{W}_m^{(2)}$ we obtain a special integral $G(\mathbb{Z})$-orbit in V whose elements have invariants b. We additionally consider a distinguished subspace $W_0(\mathbb{Z}) \subset V(\mathbb{Z})$, and we define a Q-invariant for the elements of $W_0(\mathbb{Z})$. Then, we will see that the elements in the constructed orbit have large Q-invariant when they intersect $W_0(\mathbb{Z})$ (which happens always except for a negligible amount of times by cutting-off-the-cusp arguments). This construction is the analogue of [7, Sections 2.2 and 3.2]; we give a more detailed comparison at the end of Section 3.

In view of all that, to prove Theorem 1.2 it suffices to bound the number of these distinguished $G(\mathbb{Z})$-orbits in $W_0(\mathbb{Z})$ having large Q-invariant. However, before doing that, we will need to take a small detour and estimate the number of *all* reducible $G(\mathbb{Z})$-orbits in $V(\mathbb{Z})$. A $G(\mathbb{C})$-orbit in $V(\mathbb{C})$ can split into multiple $G(\mathbb{Q})$-orbits, and among these $G(\mathbb{Q})$-orbits there is a "distinguished" one (namely, the one given by the Kostant section, as defined in Section 2.3). We say that an element in $V(\mathbb{Q})$ is *reducible* if it falls into this special $G(\mathbb{Q})$-orbit. Using Bhargava's geometry-of-numbers arguments, and in particular the techniques in the cusp developed in [31], we will obtain the following result:

Theorem 1.3. *The number $N(V(\mathbb{Z})^{red}, X)$ of reducible $G(\mathbb{Z})$-orbits on $V(\mathbb{Z})$ of height at most X is*

$$N(V(\mathbb{Z})^{red}, X) = CX^{\dim V} + O(X^{\dim V-\delta}),$$

where C, δ are real positive constants. The constant C will be explicitly determined in Section 5.5.

The proof of this theorem relies on the construction of a *box-shaped* fundamental domain for the action of $G(\mathbb{Z})$ on $G(\mathbb{R})$, which will be carried out in Section 4.3. Following that, we will mostly follow the steps in [31, Section 4], relying on critical reductions given by [21, Section 8]; and we will conclude our proof by using elementary but lengthy case-by-case computations. We remark that our proof implicitly also relies on other case-by-case computations: namely, the cutting-off-the-cusp result in Proposition 5.3 relies on an (even more tedious) exhaustive analysis of all cases, sometimes relying on lengthy computations on a computer (cf. [28, Proposition 4.5]). It remains open to see whether there is any uniform proof to these results.

In Section 6, we will conclude the proof of Theorem 1.2, from which Theorem 1.1 will follow using a squarefree sieve. The sieve is carried out in a general enough setting that allows us to count the density of subsets in $B(\mathbb{Z})$ defined by infinitely many congruence conditions. In particular, we get an application of our result to the context of [21], which allows us to get an upper bound on the average size of 2-Selmer groups of families defined by infinitely many congruence conditions. For $b \in B(\mathbb{Z})$, denote by J_b the Jacobian of the curve C_b.

Theorem 1.4. *Let m be the number of marked points of the family $C \to B$, as given in Table 2.2. Let $\mathcal{S}$ be a κ-acceptable subset of $B(\mathbb{Z})$ in the sense of Section 6.4. Then, we have*

$$\limsup_{X\to\infty} \frac{\sum_{b\in\mathcal{S},\, \mathrm{ht}(b)<X} \#\,\mathrm{Sel}_2\, J_b}{\#\{b \in \mathcal{S} \mid \mathrm{ht}(b) < X\}} \leq 3 \cdot 2^{m-1}.$$

Acknowledgements. This paper was written while the author was a PhD student under the supervision of Jack Thorne. I would like to thank him for providing many useful suggestions, guidance and encouragement during the process, and for revising an early version of this manuscript. I also wish to thank Jef Laga and the anonymous referees for their helpful comments.

2. Preliminaries

In this section, we introduce our representation (G, V) of interest, together with some of its basic properties. We do so mostly following [37, Section 2] and [21, Section 3].

2.1. Vinberg representations. Let H be a split adjoint simple group of type A, D, E over $\mathbb{Q}$. We assume H is equipped with a pinning $(T, P, \{X_\alpha\})$, meaning:

- $T \subset H$ is a split maximal torus (determining a root system Φ_H).

- $P \subset H$ is a Borel subgroup containing T (determining a root basis $S_H \subset \Phi_H$).
- X_α is a generator for $\mathfrak{h}_\alpha$ for each $\alpha \in S_H$.

Let $W = N_H(T)/T$ be the Weyl group of Φ_H, and let $\mathcal{D}$ be the Dynkin diagram of H. Then, we have the following exact sequences:

$$0 \longrightarrow H \longrightarrow \mathrm{Aut}(H) \longrightarrow \mathrm{Aut}(\mathcal{D}) \longrightarrow 0 \tag{2.1}$$

$$0 \longrightarrow W \longrightarrow \mathrm{Aut}(\Phi_H) \longrightarrow \mathrm{Aut}(\mathcal{D}) \longrightarrow 0 \tag{2.2}$$

The subgroup $(T, P, \{X_\alpha\}) \subset \mathrm{Aut}(H)$ of automorphisms of H preserving the pinning determines a splitting of (2.1). Then, we can define $\vartheta \in \mathrm{Aut}(H)$ as the unique element in $(T, P, \{X_\alpha\})$ such that its image in $\mathrm{Aut}(\mathcal{D})$ under (2.1) coincides with the image of $-1 \in \mathrm{Aut}(\Phi_H)$ under (2.2). Writing $\check{\rho}$ for the sum of fundamental coweights with respect to S_H, we define

$$\theta := \vartheta \circ \mathrm{Ad}(\check{\rho}(-1)) = \mathrm{Ad}(\check{\rho}(-1)) \circ \vartheta.$$

The map θ defines an involution of H, and so $d\theta$ defines an involution of the Lie algebra $\mathfrak{h}$. By considering ± 1 eigenspaces, we obtain a $\mathbb{Z}/2\mathbb{Z}$-grading

$$\mathfrak{h} = \mathfrak{h}(0) \oplus \mathfrak{h}(1),$$

where $[\mathfrak{h}(i), \mathfrak{h}(j)] \subset \mathfrak{h}(i+j)$. We define $G = (H^\theta)^\circ$ and $V = \mathfrak{h}(1)$, which means that V is a representation of G by restriction of the adjoint representation. Moreover, we have $\mathrm{Lie}(G) = \mathfrak{h}(0)$.

We have the following basic result [24, Theorem 1.1] on the GIT quotient $B := V /\!\!/ G = \mathrm{Spec}\, \mathbb{Q}[V]^G$.

Theorem 2.1. *Let $\mathfrak{c} \subset V$ be a Cartan subspace. Then, $\mathfrak{c}$ is a Cartan subalgebra of $\mathfrak{h}$, and the map $N_G(\mathfrak{c}) \to W_\mathfrak{c} := N_H(\mathfrak{c})/Z_H(\mathfrak{c})$ is surjective. Therefore, the canonical inclusions $\mathfrak{c} \subset V \subset \mathfrak{h}$ induce isomorphisms*

$$\mathfrak{c} /\!\!/ W_\mathfrak{c} \cong V /\!\!/ G \cong \mathfrak{h} /\!\!/ H.$$

In particular, all these quotients are isomorphic to a finite-dimensional affine space.

For any field k of characteristic zero, we can define the *discriminant polynomial* $\Delta \in k[\mathfrak{h}]^H$ as the image of $\prod_{\alpha \in \Phi_H} \alpha$ under the isomorphism $k[\mathfrak{t}]^W \xrightarrow{\sim} k[\mathfrak{h}]^H$. The discriminant can also be regarded as a polynomial in $k[B]$ through the isomorphism $k[\mathfrak{h}]^H \cong k[V]^G = k[B]$. We can relate the discriminant to one-parameter subgroups, which we now introduce. If $k/\mathbb{Q}$ is a field and $\lambda \colon \mathbb{G}_m \to G_k$ is a homomorphism, there exists a decomposition $V = \sum_{i \in \mathbb{Z}} V_i$, where $V_i := \{v \in V(k) \mid \lambda(t)v = t^i v \ \ \forall\, t \in \mathbb{G}_m(k)\}$. Every vector $v \in V(k)$ can be written as $v = \sum v_i$, where $v_i \in V_i$; we call the integers i with $v_i \neq 0$ the *weights* of v. Finally, we recall that an element $v \in \mathfrak{h}$ is *regular* if its centraliser has minimal dimension.

Proposition 2.2. *Let $k/\mathbb{Q}$ be a field, and let $v \in V(k)$. The following are equivalent:*

(1) *v is regular semisimple.*
(2) *$\Delta(v) \neq 0$.*
(3) *For every non-trivial homomorphism $\lambda\colon \mathbb{G}_m \to G_{k^s}$, v has a positive weight with respect to λ.*

Proof. The reasoning is the same as in [28, Corollary 2.4]. □

We remark that the Vinberg representation (G, V) can be identified explicitly. For the reader's convenience, we reproduce the explicit description written in [21, Section 3.2] in Table 2.1. We refer the reader to loc. cit. for the precise meaning of some of these symbols.

TABLE 2.1. Explicit description of each representation

Type	G	V
A_{2n}	SO_{2n+1}	$\mathrm{Sym}^2(2n+1)_0$
A_{2n+1}	PSO_{2n+2}	$\mathrm{Sym}^2(2n+2)_0$
D_{2n} $(n \geq 2)$	$\mathrm{SO}_{2n} \times \mathrm{SO}_{2n} / \Delta(\mu_2)$	$2n \boxtimes 2n$
D_{2n+1} $(n \geq 2)$	$\mathrm{SO}_{2n+1} \times \mathrm{SO}_{2n+1}$	$(2n+1) \boxtimes (2n+1)$
E_6	PSp_8	$\wedge^4_0 8$
E_7	SL_8 / μ_4	$\wedge^4 8$
E_8	Spin_{16}/μ_2	half spin

2.2. Restricted roots. In the previous section, we considered the root system $\Phi := \Phi_H$ of H, but we will also need to understand the restricted root system $\Phi(G, T^\theta)$ and the set of weights Φ_V of the action of T^θ on V. This will be particularly important when defining the distinguished subspace $W_0 \subset V$ and the Q-invariant in Section 3. The exposition in this section is based on [38, Section 2.3].

Write Φ/ϑ for the orbits of ϑ on Φ, where ϑ is the pinned automorphism defined in the previous section.

Lemma 2.3.

(1) *The map $X^*(T) \to X^*(T^\theta)$ is surjective, and the group G is adjoint. In particular, $X^*(T^\theta)$ is spanned by $\Phi(G, T^\theta)$.*
(2) *Let $\alpha, \beta \in \Phi$. Then, the image of α in $X^*(T^\theta)$ is non-zero, and α, β have the same image if and only if either $\alpha = \beta$ or $\alpha = \vartheta(\beta)$.*

Proof. This is [38, Lemma 2.5]. □

Hence, we can identify Φ/ϑ with its image in $X^*(T^\theta)$. We note that $\vartheta = 1$ if and only if -1 is an element of the Weyl group $W(H, T)$; in this case Φ/ϑ coincides with Φ.

We can write the following decomposition:

$$\mathfrak{h} = \mathfrak{t} \oplus \bigoplus_{a \in \Phi/\vartheta} \mathfrak{h}_a,$$

with $\mathfrak{t} = \mathfrak{t}^\theta \oplus V_0$ and $\mathfrak{h}_a = \mathfrak{g}_a \oplus V_a$, so that

$$\mathfrak{g} = \mathfrak{t}^\theta \bigoplus_{a \in \Phi/\vartheta} \mathfrak{g}_a, \quad V = V_0 \oplus \bigoplus_{a \in \Phi/\vartheta} V_a.$$

Given $a \in \Phi/\vartheta$, we can identify $\mathfrak{g}_a$ and V_a explicitly according to the value of $s = (-1)^{\langle \alpha, \check{\rho} \rangle}$:

(1) $a = \{\alpha\}$ and $s = 1$. Then, $V_a = 0$ and $\mathfrak{g}_\alpha$ is spanned by X_α.
(2) $a = \{\alpha\}$ and $s = -1$. Then, V_a is spanned by X_α and $\mathfrak{g}_\alpha = 0$.
(3) $a = \{\alpha, \vartheta(\alpha)\}$, with $\alpha \neq \vartheta(\alpha)$. Then, V_a is spanned by $X_\alpha - sX_{\vartheta(\alpha)}$ and $\mathfrak{g}_\alpha$ is spanned by $X_\alpha + sX_{\vartheta(\alpha)}$.

We note that ϑ preserves the height of a root α with respect to the basis S_H (recall that the height of a root α is defined as $\sum_i c_i$, where $\alpha = \sum_{\alpha_i \in S_H} c_i \alpha_i$ is the decomposition as the sum of simple roots). Therefore, it will make sense to define the *height* of a root $a \in \Phi/\vartheta$ as the height of any element in $\vartheta^{-1}(a)$.

Remark 2.4. It will be important for us to define the height of a root in Φ/ϑ relative to its corresponding height in $\mathfrak{h}$ and not relative to its height with respect to some choice of basis of the root system $\Phi(G, T^\theta)$. As an example, consider the E_6 case following the conventions of [38]. Say that a root basis for Φ_H is $\{\alpha_1, \ldots, \alpha_6\}$ and a basis for $\Phi(G, T^\theta)$ is $(a_1, a_2, a_3, a_4) = (\alpha_3 + \alpha_4, \alpha_1, \alpha_3, \alpha_2 + \alpha_4)$. If a root $a \in \Phi_V$ can be expressed as $a = \sum n_i a_i = \sum m_i \alpha_i$ for some integers $n_i, m_i \in \mathbb{Z}$, under our definitions the height of a root is $\mathrm{ht}(\alpha) = \sum m_i$ and not $\sum n_i$. This is different than the natural notion of height we might arrive at if we consider the weights of V as a representation of G abstractly.

2.3. Transverse slices over $V /\!\!/ G$. In this section, we present some remarkable properties of the map $\pi \colon V \to B$, where we recall that $B := V /\!\!/ G$ is the GIT quotient.

Definition 2.5. An $\mathfrak{sl}_2$*-triple* of $\mathfrak{h}$ is a triple (e, h, f) of non-zero elements of $\mathfrak{h}$ satisfying

$$[h, e] = 2e, \quad [h, f] = -2f, \quad [e, f] = h.$$

Moreover, we say this $\mathfrak{sl}_2$-triple is *normal* if $e, f \in \mathfrak{h}(1)$ and $h \in \mathfrak{h}(0)$.

Theorem 2.6 (Graded Jacobson–Morozov)**.** *Every non-zero nilpotent element $e \in \mathfrak{h}(1)$ is contained in a normal $\mathfrak{sl}_2$-triple. If e is also regular, then it is contained in a unique normal $\mathfrak{sl}_2$-triple.*

TABLE 2.2. Families of curves

Type	Curve	Marked points
A_{2n}	$y^2 = x^{2n+1} + p_2x^{2n-1} + \cdots + p_{2n+1}$	1
A_{2n+1}	$y^2 = x^{2n+2} + p_2x^{2n} + \cdots + p_{2n+1}$	2
D_{2n}	$y(xy + p_{2n}) = x^{2n-1} + p_2x^{2n-2} + \cdots + p_{4n-2}$	3
D_{2n+1}	$y(xy + p_{2n+1}) = x^{2n} + p_2x^{2n-1} + \cdots + p_{4n}$	2
E_6	$y^3 = x^4 + (p_2x^2 + p_5x + p_8)y + (p_6x^2 + p_9x + p_{12})$	1
E_7	$y^3 = x^3y + p_{10}x^2 + x(p_2y^2 + p_8y + p_{14}) + p_6y^2 + p_{12}y + p_{18}$	2
E_8	$y^3 = x^5 + (p_2x^3 + p_8x^2 + p_{14}x + p_{20})y + (p_{12}x^3 + p_{18}x^2 + p_{24}x + p_{30})$	1

Proof. The first part of the statement is [37, Lemma 2.17], and the second part follows from [37, Lemma 2.14]. □

Definition 2.7. Let r be the rank of $\mathfrak{h}$. We say an element $x \in \mathfrak{h}$ is *subregular* if $\dim \mathfrak{z}_{\mathfrak{h}}(x) = r + 2$.

Subregular nilpotent elements in V exist by [37, Proposition 2.27]. Let $e \in V$ be such an element, and fix a normal $\mathfrak{sl}_2$-triple (e, h, f) using Theorem 2.6. Let $C = e + \mathfrak{z}_V(f)$, and consider the natural morphism $\varphi\colon C \to B$.

Theorem 2.8.

(1) *The geometric fibres of the morphism φ are reduced connected curves. For $b \in B(k)$, the corresponding curve C_b is smooth if and only if $\Delta(b) \neq 0$.*
(2) *The central fibre $\varphi^{-1}(0)$ has a unique singular point which is a simple singularity of type A_n, D_n, E_n, coinciding with the type of H.*
(3) *We can choose coordinates $p_{d_1}, \ldots, p_{d_r}$ in B, with p_{d_i} being homogeneous of degree d_i, and coordinates $(x, y, p_{d_1}, \ldots, p_{d_r})$ on C such that $C \to B$ is given by Table 2.2.*

Proof. See [37, Theorem 3.8]. □

Our choice of pinning in Section 2.1 determines a natural choice of a regular nilpotent element, namely $E = \sum_{\alpha \in S_H} X_\alpha \in V(\mathbb{Q})$. Let (E, H, F) be its associated normal $\mathfrak{sl}_2$-triple by Theorem 2.6. We define the affine linear subspace $\kappa_E := (E + \mathfrak{z}_{\mathfrak{h}}(F)) \cap V$ as the *Kostant section* associated to E. Whenever E is understood, we will just denote the Kostant section by κ.

Theorem 2.9. *The composition $\kappa \hookrightarrow V \to B$ is an isomorphism, and every element of κ is regular.*

Proof. See [37, Lemma 3.5]. □

Definition 2.10. Let $k/\mathbb{Q}$ be a field and let $v \in V(k)$. We say v is *k-reducible* if $\Delta(v) = 0$ or if v is $G(k)$-conjugate to some Kostant section, and *k-irreducible* otherwise.

We will typically refer to $\mathbb{Q}$-(ir)reducible elements simply as (ir)reducible. We note that if k is algebraically closed, then all elements of V are reducible, see [21, Proposition 2.11].

2.4. Integral structures. So far, we have considered our objects of interest over $\mathbb{Q}$, but for our purposes it will be crucial to define integral structures for G and V.

The structure of G over $\mathbb{Z}$ comes from the general classification of split reductive groups over any non-empty scheme S: namely, every root datum is isomorphic to the root datum of a split reductive S-group (see [12, Theorem 6.1.16]). By considering the root datum $\Phi(G, T^\theta)$ studied in Section 2.2 and the scheme $S = \operatorname{Spec}\mathbb{Z}$, we get a split reductive group $\underline{G}$ defined over $\mathbb{Z}$, such that its base change to $\mathbb{Q}$ coincides with G. By [27, Lemma 5.1], we know that T^θ, P^θ are a maximal split torus and a Borel subgroup of G, respectively. We also get integral structures for $\underline{T}^\theta$ and $\underline{P}^\theta$ inside of $\underline{G}$.

Proposition 2.11. *$\underline{G}$ and $\underline{P}^\theta$ have class number 1: $\underline{G}(\mathbb{A}^\infty) = G(\mathbb{Q})\underline{G}(\widehat{\mathbb{Z}})$ and $\underline{P}^\theta(\mathbb{A}^\infty) = P^\theta(\mathbb{Q})\underline{P}^\theta(\widehat{\mathbb{Z}})$.*

Proof. Note that $\mathrm{cl}(\underline{G}) \leq \mathrm{cl}(\underline{P}^\theta) \leq \mathrm{cl}(\underline{T}^\theta)$ by [25, Theorem 8.11, Corollary 1]. We see that $\underline{T}$ has class number 1 using [25, Theorem 8.11, Corollary 2], since G contains a $\mathbb{Q}$-split torus consisting of diagonal matrices in $\mathrm{GL}(V)$ and $\mathbb{Q}$ has class number 1. □

To obtain the $\mathbb{Z}$-structure for V, we consider $\mathfrak{h}$ as a semisimple G-module over $\mathbb{Q}$ via the restriction of the adjoint representation. This G-module splits into a sum of simple G-modules:

$$\mathfrak{h} = \left(\bigoplus_{i=1}^{r} V_i\right) \oplus \left(\bigoplus_{i=1}^{s} \mathfrak{g}_i\right),$$

where $\bigoplus V_i = V$ and $\bigoplus \mathfrak{g}_i = \mathfrak{g}$, since both subspaces are G-invariant. For each of these irreducible representations, we can choose highest weight vectors $v_i \in V_i$ and $w_i \in \mathfrak{g}_i$, and we then consider

$$\underline{V}_i := \mathrm{Dist}(\underline{G})v_i, \quad \underline{\mathfrak{g}}_i := \mathrm{Dist}(\underline{G})w_i,$$

where $\mathrm{Dist}(\underline{G})$ the algebra of distributions of $\underline{G}$ (see [18, I.7.7]). By the results in [18, II.8.3], we have that $V_i = \mathbb{Q} \otimes_{\mathbb{Z}} \underline{V}_i$, $\mathfrak{g}_i = \mathbb{Q} \otimes_{\mathbb{Z}} \underline{\mathfrak{g}}_i$ and that $\underline{V} := \bigoplus \underline{V}_i$ is a $\underline{G}$-stable lattice inside V. By scaling the highest weight vectors if necessary, we will assume that $E \in \underline{V}(\mathbb{Z})$.

We can also consider an integral structure $\underline{B}$ on B. We can take the polynomials $p_{d_1}, \ldots, p_{d_r} \in \mathbb{Q}[V]^G$ determined in Section 2.3 and rescale

them using the $\mathbb{G}_m$-action $t \cdot p_{d_i} = t^{d_i} p_{d_i}$ to make them lie in $\mathbb{Z}[\underline{V}]^{\underline{G}}$. We let $\underline{B} := \operatorname{Spec} \mathbb{Z}[p_{d_1}, \dots, p_{d_r}]$ and write $\pi \colon \underline{V} \to \underline{B}$ for the corresponding morphism. We may additionally assume that the discriminant Δ defined in Section 2.1 lies in $\mathbb{Z}[\underline{V}]^{\underline{G}}$, where the coefficients of Δ in $\mathbb{Z}[p_{d_1}, \dots, p_{d_r}]$ may be assumed to not have a common divisor.

A crucial step in our argument will be to make our constructions in $\mathbb{Z}_p$ for all p and then glue it all together using the class number one property in Proposition 2.11. For this, we will need the following lemma, which records the existence of orbits in $\underline{V}(\mathbb{Z}_p)$ (cf. [38, Lemma 2.8]):

Lemma 2.12. *There exists an integer $N_0 \geq 1$ such that for all primes p and for all $b \in \underline{B}(\mathbb{Z}_p)$ we have $N_0 \cdot \kappa_b \in \underline{V}(\mathbb{Z}_p)$.*

Our arguments in Section 3 will implicitly rely on integral geometric properties of the representation (G, V). In there, we will need to avoid finitely many primes, or more precisely to work over $S = \operatorname{Spec} \mathbb{Z}[1/N]$ for a suitable $N \geq 1$. By combining the previous lemma and the spreading out properties in [21, Section 7.2], we get:

Proposition 2.13. *There exists a positive integer $N \geq 1$ such that:*

1. *For every $b \in \underline{B}(\mathbb{Z})$, the corresponding Kostant section κ_b is $G(\mathbb{Q})$-conjugate to an element in $\frac{1}{N}\underline{V}(\mathbb{Z})$.*
2. *N is admissible in the sense of* [21, Section 7.2].

In particular, we will always assume that N is even. We fix the integer N in Proposition 2.13 throughout the rest of the paper. We will also drop the underline notation for the objects defined over $\mathbb{Z}$, and just refer to $\underline{G}, \underline{V} \dots$ as $G, V \dots$ by abuse of notation.

To end this section, we consider some further integral properties of the Kostant section. In Section 2.3, we considered κ defined over $\mathbb{Q}$, and now we will consider some of its properties over $\mathbb{Z}_p$. Consider the decomposition

$$\mathfrak{h} = \bigoplus_{j \in \mathbb{Z}} \mathfrak{h}_j$$

according to the height of the roots. If P^- is the negative Borel subgroup of H, N^- is its unipotent radical and $\mathfrak{p}^-$ and $\mathfrak{n}^-$ are their respective Lie algebras, we have $\mathfrak{p}^- = \bigoplus_{j \leq 0} \mathfrak{h}_j$, $\mathfrak{n}^- = \bigoplus_{j<0} \mathfrak{h}_j$ and $[E, \mathfrak{h}_j] \subset \mathfrak{h}_{j+1}$.

Theorem 2.14. *Let R be a ring in which N is invertible. Then:*

1. *$[E, \mathfrak{n}_R^-]$ has a complement in $\mathfrak{p}_R^-$ of rank $\operatorname{rk}_R \mathfrak{p}_R^- - \operatorname{rk}_R \mathfrak{n}_R^-$; call it Ξ.*
2. *The action map $N^- \times (E + \Xi) \to E + \mathfrak{p}^-$ is an isomorphism over R.*
3. *Both maps in the composition $E + \Xi \to (E + \mathfrak{p}^-) /\!\!/ N^- \to \mathfrak{h} /\!\!/ H$ are isomorphisms over R.*

Proof. See [1, Section 2.3]. □

Remark 2.15. If R is a field of characteristic not dividing N, then Ξ can be taken to be $\mathfrak{z}_{\mathfrak{h}}(F)$ and $E+\Xi$ is the same as the Kostant section considered in Section 2.3. We will abuse notation by referring to both the Kostant section defined in Section 2.3 and the section in Theorem 2.14 by κ.

Theorem 2.14 will be an important improvement from Theorem 2.9, since in the sequel we will need the Kostant section to maintain certain integrality properties. In particular, it will be helpful to apply Theorem 2.14 over $\mathbb{Z}_p$, a feature that would not be present if we only had Theorem 2.9.

3. Constructing orbits

Given an element $b \in B(\mathbb{Z})$ with discriminant weakly divisible by m^2 for a large squarefree number m coprime to N, we will show how to construct a special $g \in G(\mathbb{Z}[1/m]) \setminus G(\mathbb{Z})$ such that $g\kappa_b \in \frac{1}{N}V(\mathbb{Z})$ in a way that "remembers m".

We start by defining the distinguished subspace $W_0 \subset V$ as

$$W_0 := \bigoplus_{\substack{a \in \Phi/\vartheta \\ \mathrm{ht}(a) \leq 1}} V_a,$$

where the notation is as in Section 2.2. We write an element $v \in W_0(\mathbb{Q})$ as $v = \sum_{\mathrm{ht}(\alpha)=1} v_\alpha X_\alpha + \sum_{\mathrm{ht}(\beta)\leq 0} v_\beta X_\beta$, where each X_α, X_β generates each root space V_α, V_β and $v_\alpha, v_\beta \in \mathbb{Q}$. Then, we can define the *Q-invariant* of $v \in W_0$ as $Q(v) = \left|\prod_{\mathrm{ht}(\alpha)=1} v_\alpha\right|$. Now, define:

$$W_M := \left\{ v \in \frac{1}{N}V(\mathbb{Z}) \;\middle|\; \begin{array}{l} v = g\kappa_b \text{ for a squarefree } m > M,\ (m,N)=1, \\ g \in G(\mathbb{Z}[1/m]) \setminus G(\mathbb{Z}),\ b \in B(\mathbb{Z}), \Delta(b) \neq 0 \end{array} \right\}.$$

The main result of the section is the following:

Proposition 3.1. *Let $b \in B(\mathbb{Z})$, and assume that $\mathrm{Stab}_{G(\mathbb{Q})} \kappa_b = \{e\}$.*

1. *Let $m > M$ be a squarefree integer, coprime to N. If m^2 weakly divides $\Delta(b)$, then $W_M \cap \pi^{-1}(b)$ is non-empty.*
2. *If $v \in W_M \cap W_0$, then $Q(v) > M$.*

The proof of Proposition 3.1 will rely on a reduction to $\mathfrak{sl}_2$, inspired by the techniques in the proofs of [20, Lemma 4.19] and [29, Proposition 5.4], which we now explain.

Assume we have a connected reductive group L over a field k, together with an involution ξ. As in Section 2.1, the Lie algebra $\mathfrak{l}$ decomposes as $\mathfrak{l} = \mathfrak{l}(0) \oplus \mathfrak{l}(1)$, according to the ± 1 eigenspaces of $d\xi$. We also write L_0 for the connected component of the fixed group L^ξ.

Definition 3.2. Let k be algebraically closed. We say a vector $v \in \mathfrak{l}(1)$ is *stable* if the L_0-orbit of v is closed and its stabiliser $Z_{L_0}(v)$ is finite. We

say $(L_0, \mathfrak{l}(1))$ is *stable* if it contains stable vectors. If k is not necessarily algebraically closed, we say $(L_0, \mathfrak{l}(1))$ is *stable* if $(L_{0,k^s}, \mathfrak{l}(1)_{k^s})$ is.

By [39, Proposition 1.9], the θ defined in Section 2.1 is a stable involution, i.e. (G, V) is stable.

We now prove the analogue of [29, Lemma 2.3]: the proof is very similar and is reproduced for convenience.

Lemma 3.3. *Let S be a $\mathbb{Z}[1/N]$-scheme. Let (L, ξ), (L', ξ') be two pairs, each consisting of a reductive group over S whose geometric fibres are adjoint semisimple of type A_1, together with a stable involution. Then for any $s \in S$ there exists an étale morphism $S' \to S$ with image containing s and an isomorphism $L_{S'} \to L_S$ intertwining $\xi_{S'}$ and $\xi'_{S'}$.*

Proof. We are working étale locally on S, so we can assume that $L = L'$ and that they are both split reductive groups. Let T denote the scheme of elements $l \in L$ such that $\mathrm{Ad}(l) \circ \xi = \xi'$: by [12, Proposition 2.1.2], T is a closed subscheme of L that is smooth over S. Since a surjective smooth morphism has sections étale locally, it is sufficient to show that $T \to S$ is surjective. Moreover, we can assume that $S = \operatorname{Spec} k$ for an algebraically closed field k, since the formation of T commutes with base change.

Let $A, A' \subset L$ be maximal tori on which ξ, ξ' act as an automorphism of order 2. By the conjugacy of maximal tori, we can assume that $A = A'$ and that ξ, ξ' define the (unique) element of order 2 in the Weyl group. Write $\xi = a\xi'$ for some $a \in A(k)$. Writing $a = b^2$ for some $b \in A(k)$, we have $\xi = b \cdot b \cdot \xi' = b \cdot \xi' \cdot b^{-1}$. The conclusion is that ξ and ξ' are $L(k)$-conjugate (in fact, $A(k)$-conjugate), which completes the proof. □

The following lemma is the key technical part in our proof. We remark the first part was already implicitly proven in the proof of [21, Theorem 7.17].

Lemma 3.4. *Let p be a prime that does not divide N.*

1. *Let $b \in B(\mathbb{Z}_p)$ be an element with $\mathrm{ord}_p \Delta(b) = 1$, where $\mathrm{ord}_p \colon \mathbb{Q}_p^* \to \mathbb{Z}$ is the usual normalised valuation. Let $v \in V(\mathbb{Z}_p)$ with $\pi(v) = b$. Then, the reduction mod p of v in $V(\mathbb{F}_p)$ is regular.*
2. *Let $b \in B(\mathbb{Z}_p)$ be an element with discriminant weakly divisible by p^2. Then, there exists $g_{b,p} \in G(\mathbb{Q}_p) \setminus G(\mathbb{Z}_p)$ such that $g_{b,p} \cdot \kappa_b \in V(\mathbb{Z}_p)$.*

Proof. Let $v_{\mathbb{F}_p} = x_s + x_n$ be the Jordan decomposition of the reduction of v in $\mathbb{F}_p$. Then, we have a decomposition $\mathfrak{h}_{\mathbb{F}_p} = \mathfrak{h}_{0,\mathbb{F}_p} \oplus \mathfrak{h}_{1,\mathbb{F}_p}$, where $\mathfrak{h}_{0,\mathbb{F}_p} = \mathfrak{z}_{\mathfrak{h}}(x_s)$ and $\mathfrak{h}_{1,\mathbb{F}_p} = \mathrm{image}(\mathrm{Ad}(x_s))$. By Hensel's lemma, this decomposition lifts to $\mathfrak{h}_{\mathbb{Z}_p} = \mathfrak{h}_{0,\mathbb{Z}_p} \oplus \mathfrak{h}_{1,\mathbb{Z}_p}$, with $\mathrm{ad}(v)$ acting topologically nilpotently in $\mathfrak{h}_{0,\mathbb{Z}_p}$ and invertibly in $\mathfrak{h}_{1,\mathbb{Z}_p}$. As explained in the proof of [20, Lemma 4.19], there is a unique closed subgroup $L \subset H_{\mathbb{Z}_p}$ which is smooth over $\mathbb{Z}_p$ with connected fibres and with Lie algebra $\mathfrak{h}_{0,\mathbb{Z}_p}$.

For the first part of the lemma, we are free to replace $\mathbb{Z}_p$ for a complete discrete valuation ring R with uniformiser p, containing $\mathbb{Z}_p$ and with algebraically closed residue field k. In this case, the spreading out properties in [21, Section 7.2] guarantee that the derived group of L is of type A_1. Since the restriction of θ restricts to a stable involution in L by [37, Lemma 2.5], Lemma 3.3 guarantees that there exists an isomorphism $\mathfrak{h}_{0,R}^{der} \cong \mathfrak{sl}_{2,R}$ intertwining the action of θ on $\mathfrak{h}_{0,R}^{der}$ with the action of $\xi = \mathrm{Ad}(\mathrm{diag}(1,-1))$ on $\mathfrak{sl}_{2,R}$. To show that v_k is regular is equivalent to showing that the nilpotent part x_n is regular in $\mathfrak{h}_{0,k}^{der}$. The elements v_k and x_n have the same projection in $\mathfrak{h}_{0,k}^{der}$, and given that $v \in \mathfrak{h}_{0,R}^{der,d\theta=-1}$, its image in $\mathfrak{sl}_{2,R}$ is of the form

$$\begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}.$$

We claim that $\mathrm{ord}_R(ab) = 1$. This can be seen from an argument similar to the end of [21, Lemma 7.15], i.e. using [21, Lemma 2.3][1], it follows that the discriminant of v in $\mathfrak{h}$ coincides with the discriminant of its image in $\mathfrak{sl}_2$ up to a unit in R, as wanted. In particular, exactly one of a, b is non-zero when reduced to k, and hence x_n is regular in $\mathfrak{h}_{0,k}^{der}$, as wanted.

For the second part, we return to the case $R = \mathbb{Z}_p$. If $b \in B(\mathbb{Z}_p)$ has discriminant weakly divisible by p^2, there exists $b' \in B(\mathbb{Z}_p)$ such that $\mathrm{ord}_p \Delta(b + pb') = 1$. Since the Kostant section κ is algebraic, we know that $\kappa_b - \kappa_{b+pb'} \in pV(\mathbb{Z}_p)$. By the first part of the lemma, we know that $\kappa_{b+pb'}$ is regular mod p, and so κ_b is also regular mod p. In particular, writing $\kappa_{b,\mathbb{F}_p} = x_s + x_n$ as before, this means that the nilpotent part x_n is a regular nilpotent in $\mathfrak{h}_{0,\mathbb{F}_p}^{der}$. We now claim that:

(1) We have an isomorphism $\mathfrak{h}_{0,\mathbb{Z}_p}^{der} \cong \mathfrak{sl}_{2,\mathbb{Z}_p}$;
(2) The isomorphism intertwines the actions of θ and the previously defined ξ;
(3) Over $\mathbb{F}_p$, the isomorphism sends the regular nilpotent x_n to the matrix
$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$
of $\mathfrak{sl}_{2,\mathbb{F}_p}$.

We note that this does not follow immediately from Lemma 3.3, as the isomorphism a priori does not need to be defined over $\mathbb{Z}_p$.

We prove our claim as follows: Let $X = \mathrm{Isom}((L/Z(L), \theta), (\mathrm{PGL}_2, \xi))$ be the scheme over $\mathbb{Z}_p$ consisting of isomorphisms between $L/Z(L)$ and PGL_2 that intertwine the θ and ξ-actions. Using Lemma 3.3, we see that

[1]The cited result is only stated in [21] for fields of characteristic zero, but it is still valid in our situation for k: the only results that are invoked in that proof are those in [35, Section 3], which hold as long as $\mathrm{char}\, k$ is not a torsion prime for H, which we may assume.

étale-locally, X is isomorphic to $\mathrm{Aut}(\mathrm{PGL}_2, \xi)$; in particular, it is a smooth scheme over $\mathbb{Z}_p$. By Hensel's lemma [15, Théorème 18.5.17], to show that X has a $\mathbb{Z}_p$-point it is sufficient to show that it has an $\mathbb{F}_p$-point.

We now look at the $\mathbb{F}_p$-scheme $Y = \mathrm{Isom}((L/Z(L)_{\mathbb{F}_p}, \theta, x_n), (\mathrm{PGL}_2, \xi, e))$ of isomorphisms preserving the θ and ξ-actions which send x_n to e: it is a subscheme of $X_{\mathbb{F}_p}$. If $Y(\mathbb{F}_p)$ is non-empty, then by Hensel's lemma it can be lifted to an isomorphism of $X(\mathbb{Z}_p)$ satisfying all three points of the claim. Therefore, the claim will follow from seeing that $Y(\mathbb{F}_p) \neq \emptyset$.

Again by Lemma 3.3, Y is étale locally of the form $\mathrm{Aut}(\mathrm{PGL}_2, \xi, e)$, since PGL_2^{ξ} acts transitively on the regular nilpotents of $\mathfrak{sl}_2^{d\xi=-1}$ for any field of characteristic $p > N$. In particular, we see that Y is an $\mathrm{Aut}(\mathrm{PGL}_2, \xi, e)$-torsor. In this situation, to see that $Y(\mathbb{F}_p)$ is non-empty it will suffice to see that $\mathrm{Aut}(\mathrm{PGL}_2, \xi, e) = \operatorname{Spec} \mathbb{F}_p$. This follows from the elementary computation of the stabiliser of e under PGL_2^{ξ}, which can be seen to be trivial over any field.

In conclusion, $Y(\mathbb{F}_p)$ is non-empty, meaning that there is an isomorphism $\mathfrak{h}^{der}_{0,\mathbb{Z}_p} \cong \mathfrak{sl}_{2,\mathbb{Z}_p}$ respecting θ and ξ, and we can make it so that the projection of κ_b in $\mathfrak{sl}_{2,\mathbb{Z}_p}$ is an element of the form

$$\begin{pmatrix} 0 & a \\ bp^2 & 0 \end{pmatrix},$$

with $a, b \in \mathbb{Z}_p$ and $a \in 1+p\mathbb{Z}_p$. Moreover, there exists a morphism $\varphi\colon \mathrm{SL}_2 \to L^{der}_{\mathbb{Q}_p}$ inducing the given isomorphism $\mathfrak{h}^{der}_{0,\mathbb{Q}_p} \cong \mathfrak{sl}_{2,\mathbb{Q}_p}$, since SL_2 is simply connected. The morphism φ necessarily respects the grading, and induces a map $\mathrm{SL}_2(\mathbb{Q}_p) \to L^{der}(\mathbb{Q}_p)$ on the $\mathbb{Q}_p$-points. Consider the matrix $g_{b,p} = \varphi(\mathrm{diag}(p, p^{-1}))$: it satisfies the conditions of the lemma, and so we are done. □

Remark 3.5. A natural follow-up question to Lemma 3.4 is to ask how many $g_{b,p} \in G(\mathbb{Q}_p) \setminus G(\mathbb{Z}_p)$ are there (up to a $G(\mathbb{Z}_p)$-action) such that $g_{b,p} \cdot k_b \in V(\mathbb{Z}_p)$. The proof of the lemma implies that if $p^k \mid \Delta(b)$, then the projection of κ_b in $\mathfrak{sl}_{2,\mathbb{Z}_p}$ is of the form $\left(\begin{smallmatrix} 0 & a \\ bp^k & 0 \end{smallmatrix}\right)$, so we can conjugate by $\mathrm{diag}(p, p^{-1})$ a total of $\lfloor \frac{k}{2} \rfloor$ times. It is natural to expect that all the possible choices of $g_{b,p}$ arise in this fashion; however, we do not know if that is true.

Remark 3.6. It would be very convenient if in the proof of Lemma 3.4 we could obtain a $g \in \mathrm{SL}_2(\mathbb{Q}_p)$ such that

$$g \begin{pmatrix} 0 & a \\ bp^2 & 0 \end{pmatrix} g^{-1} = \begin{pmatrix} 0 & ap \\ bp & 0 \end{pmatrix},$$

in order to transform the "mod p^2" divisibility into "mod p" divisibility, but unfortunately that doesn't appear to be possible in general. If that were the case, the element $v' \in V(\mathbb{Z}_p)$ corresponding to the matrix $\left(\begin{smallmatrix} 0 & ap \\ bp & 0 \end{smallmatrix}\right)$

would not be regular modulo p, and in this situation we would be able to count such orbits using [2] (without needing geometry-of-numbers!). We note that this strategy is used in [29, Proof of Theorem 6.10], which works in their case because they are working over a $\mathbb{Z}/3\mathbb{Z}$-grading instead of a $\mathbb{Z}/2\mathbb{Z}$-grading.

Proof of Proposition 3.1. We start by proving the first item. Since G has class number 1 by Proposition 2.11, the natural map $G(\mathbb{Z})\backslash G(\mathbb{Z}[1/m]) \to \prod_{p|m} G(\mathbb{Z}_p)\backslash G(\mathbb{Q}_p)$ is a bijection. In Lemma 3.4, for each prime $p \mid m$, we constructed an element $g_{b,p} \in G(\mathbb{Z}_p)\backslash G(\mathbb{Q}_p)$, so all these elements together correspond to some element $g_b \in G(\mathbb{Z}[1/m]) \setminus G(\mathbb{Z})$. By construction, $g_b \cdot \kappa_b$ belongs to $(\bigcap_{p|m} V(\mathbb{Z}_p)) \cap V(\mathbb{Z}[1/m]) = V(\mathbb{Z})$.

We now prove the second item. Specifically, if $v \in W_M \cap W_0$ is given by $g\kappa_b$ for some $g \in G(\mathbb{Z}[1/m])\backslash G(\mathbb{Z})$, we will prove that $m \mid Q(v)$. It suffices to consider each prime $p \mid m$ separately, so assume that $g \in G(\mathbb{Z}[1/p]) \setminus G(\mathbb{Z})$. Since the group H is adjoint, there exists a $t \in T(\mathbb{Q})$ that makes all the height-one coefficients of $t\kappa_b$ be equal to one, and in this case we see that $t \in T^\theta(\mathbb{Q})$. By Theorem 2.14, there exists a unique $\gamma \in N^-(\mathbb{Q})$ such that $\gamma t \kappa_b = v$; by taking θ-invariants in the isomorphisms of Theorem 2.14. we see that $\gamma \in N^{-,\theta}(\mathbb{Q})$. Since the stabiliser is trivial, we see that $g = \gamma t$, or in other words that $g \in P^{-,\theta}(\mathbb{Z}[1/p]) \setminus P^{-,\theta}(\mathbb{Z})$.

Assume that $Q(v)$ is invertible in $\mathbb{Z}_p$, so that all the height-one coefficients of v are invertible. Then, there exists a $t' \in T(\mathbb{Z}_p)$ making all the height-one coefficients of $t'v$ be equal to one, and by Theorem 2.14, there exists at most one element γ' in $N^-(\mathbb{Z}_p)$ such that $\gamma' t' \kappa_b = v$. Consequently, $g \in P^{-,\theta}(\mathbb{Z}_p) \cap P^{-,\theta}(\mathbb{Z}[1/p]) = P^{-,\theta}(\mathbb{Z})$, a contradiction. In summary, we have that $p \mid Q(v)$ for all primes $p \mid m$, as wanted. □

Example 3.7. Our construction is inspired by the construction in [7, Sections 2.2 and 3.2] for the case A_n. In that case, $C \to B$ corresponds to the family of hyperelliptic curves $y^2 = f(x)$, where $f(x)$ has degree $n+1$ (there is a slight difference between this paper and [7], in that we consider $f(x)$ without an x^n term while they consider polynomials with a possibly non-zero linear term; we ignore this difference for now). The main goal of [7, Sections 2.2 and 3.2] is to construct an embedding

$$\sigma_m \colon \mathcal{W}_2^{(m)} \longrightarrow \frac{1}{4} W_0(\mathbb{Z}) \subset \frac{1}{4} V(\mathbb{Z}),$$

where $\sigma_m(f)$ has characteristic polynomial f and $Q(\sigma_m(f)) = m$.[2] By taking the usual pinning in SL_{n+1}, we see that V corresponds to the space of matrices in $\mathfrak{sl}_{n+1}$ which are symmetric across the antidiagonal, W_0 corresponds to the subspace of V where the entries above the superdiagonal are zero, and the height-one entries are precisely those in the superdiagonal

[2] In [7], the space that we denote as W_0 is denoted there by W_{00}.

(in [7], everything is "reflected vertically", so for instance V is the space of symmetric matrices across the diagonal; this makes no difference in the results). An explicit section of B can be taken to lie in $\frac{1}{4}W_0(\mathbb{Z})$: namely, if n is odd, the matrix

$$B(b_1,\dots,b_{n+1}) = \begin{pmatrix} 0 & 1 & & & & & & & \\ & 0 & \ddots & & & & & & \\ & & & 1 & & & & & \\ & & & 0 & 1 & & & & \\ & & & \frac{-b_2}{2} & -b_1 & 1 & & & \\ & & \iddots & -b_3 & \frac{-b_2}{2} & 0 & 1 & & \\ & \frac{-b_{n-2}}{2} & \iddots & \iddots & & & & \ddots & \\ \frac{-b_n}{2} & -b_{n-1} & \frac{-b_{n-2}}{2} & & & & & 0 & 1 \\ -b_{n+1} & \frac{-b_n}{2} & & & & & & & 0 \end{pmatrix}$$

can be seen to have characteristic polynomial $f(x) = x^{n+1} + b_1x^n + \cdots + b_nx + b_{n+1}$; if n is even, a similar matrix can be given. The main observation in this case is that if m^2 weakly divides $\Delta(f)$, then there exists an $l \in \mathbb{Z}$ such that $f(x+l) = x^{n+1} + p_1x^n + \cdots + mp_nx + m^2p_{n+1}$ (cf. [7, Proposition 2.2]). Then, if $D = \mathrm{diag}(m, 1, \dots, 1, m^{-1})$, we observe that the matrix

$$D(B(p_1,\dots,p_{n-1},mp_n,m^2p_{n+1}) + lI_{n+1})D^{-1}$$

is integral, has characteristic polynomial $f(x)$ and the entries in the superdiagonal are $(m, 1, \dots, 1, m)$. Thus, this matrix has Q-invariant m, as desired.

Remark 3.8. Our Q-invariant is slightly different to the Q-invariant defined in [7], which is defined in a slightly more general subspace of V. When restricting to $W_0(\mathbb{Q})$, their Q-invariant turns out to be a product of powers of the elements of the superdiagonal, whereas in our case we simply take the product of these elements. This difference does not affect the proof of Theorem 1.2, and we can also see that for both definitions the Q-invariant in the previous example is m.

4. Reduction theory

In light of the results in Section 3, to bound families of curves with non-squarefree discriminant it is sufficient to estimate the size of the $G(\mathbb{Z})$-invariant set W_M. Before we are able to obtain such an estimate, we will need to obtain a precise count of the number of reducible $G(\mathbb{Z})$-orbits in $V(\mathbb{Z})$. To do so, we will first need some results about reduction theory: most importantly, we will construct a box-shaped domain for the action of $G(\mathbb{Z})$ on $G(\mathbb{R})$, in the style of [31, Section 2.2].

4.1. Heights. Recall that $B = \operatorname{Spec} \mathbb{Z}[p_{d_1}, \dots, p_{d_r}]$. For any $b \in B(\mathbb{R})$, we define the *height* of b to be

$$\operatorname{ht}(b) = \sup_{i=1,\dots,r} |p_{d_i}(b)|^{1/d_i}.$$

Similarly, for every $v \in V(\mathbb{R})$ we define $\operatorname{ht}(v) := \operatorname{ht}(\pi(v))$. We record the following fact from [21, Lemma 8.1], which in particular means that the number of elements of $B(\mathbb{Z})_{<X} := \{b \in B(\mathbb{Z}) \mid \operatorname{ht}(b) < X\}$ is of order $X^{\dim V}$:

Lemma 4.1. *We have* $d_1 + \cdots + d_r = \dim_{\mathbb{Q}} V$.

4.2. Measures on G. Let $\Phi_G = \Phi(G, T^\theta)$ be the set of roots of G. The Borel subgroup P^θ of G determines a root basis S_G and a set of positive/negative roots $\Phi_G^\pm$, compatible with the choice of positive roots in H determined by the pinning of Section 2.1. Let $\overline{N}$ be the unipotent radical of the negative Borel subgroup $P^{-,\theta}$. Then, there exists a maximal compact subgroup $K \subset G(\mathbb{R})$ such that

$$\overline{N}(\mathbb{R}) \times T^\theta(\mathbb{R})^\circ \times K \longrightarrow G(\mathbb{R})$$

given by $(n, t, k) \mapsto ntk$ is a diffeomorphism; see [22, Chapter 3, Section 1]. We can choose K to be "compatible" with T; that is, we can choose a Cartan involution τ such that the fixed points of G with respect to τ is exactly K, and satisfying that $\tau|_T$ is just the inversion map. The following result is a well-known property of the Iwasawa decomposition:

Lemma 4.2. *Let dn, dt, dk be Haar measures on $\overline{N}(\mathbb{R}), T^\theta(\mathbb{R})^\circ, K$, respectively. Then, the assignment*

$$f \longmapsto \int_{n \in \overline{N}(\mathbb{R})} \int_{t \in T^\theta(\mathbb{R})^\circ} \int_{k \in K} f(ntk)\delta(t)^{-1} \mathrm{d}n \, \mathrm{d}t \, \mathrm{d}k$$

defines a Haar measure on $G(\mathbb{R})$. Here, we have that $\delta(t) = \prod_{\beta \in \Phi_G^-} \beta(t) = \det \operatorname{Ad}(t)|_{\operatorname{Lie} \overline{N}(\mathbb{R})}$.

We get the measure on $T^\theta(\mathbb{R})^\circ$ by pulling it back from the isomorphism $\prod_{\beta \in S_G} \beta \colon T^\theta(\mathbb{R})^\circ \to \mathbb{R}_{>0}^{\#S_G}$, where $\mathbb{R}_{>0}$ is given the standard Haar measure $\mathrm{d}^\times \lambda = \mathrm{d}\lambda/\lambda$. We will choose the normalisations for $\mathrm{d}n$ and $\mathrm{d}k$ in Section 5.1 in a way that will be convenient for us.

4.3. Fundamental domains. In this section, we construct a fundamental domain for the action of $G(\mathbb{Z})$ on $G(\mathbb{R})$. In view of [31], it will be useful to construct a "box-shaped" fundamental domain $\mathcal{F}$, which we will now define. For any $c > 0$, define $T_c = \{t \in T^\theta(\mathbb{R})^\circ \mid \forall\ \alpha \in S_G,\ \alpha(t) \geq c\}$. We define a *Siegel set* to be a set of the form $\mathcal{S} = \omega \cdot T_c \cdot K$, where $\omega \subset \overline{N}(\mathbb{R})$ is a compact subset, c is a positive real constant and K is the maximal compact subset fixed in Section 4.2. Then, we say that a fundamental domain $\mathcal{F}$ for

the action of $G(\mathbb{Z})$ on $G(\mathbb{R})$ is *box-shaped at infinity* if there exist two Siegel sets $\mathcal{S}_1 \subset \mathcal{F} \subset \mathcal{S}_2$ satisfying that:

(1) There exists an open subset $\mathcal{U}_1 \subset \mathcal{S}_1$ of full measure such that every $G(\mathbb{Z})$-orbit in $G(\mathbb{R})$ intersects $\mathcal{U}_1$ at most once.
(2) Every $G(\mathbb{Z})$-orbit in $G(\mathbb{R})$ intersects $\mathcal{S}_2$ at least once.
(3) For sufficiently large c, we have $\mathcal{S}_1 \cap \overline{N}T_cK = \mathcal{S}_2 \cap \overline{N}T_cK$.

To construct $\mathcal{F}$, we will see that it is sufficient to construct $\mathcal{S}_1$ and $\mathcal{S}_2$. More precisely, we have as in [31, Lemma 7]:

Lemma 4.3. *Let Λ be a discrete subgroup of a Lie group G and denote by $B(G)$ the Borel σ-algebra of G. Assume there exist sets $\mathcal{S}_1, \mathcal{S}_2$ in $B(G)$ such that the natural maps $\mathcal{S}_1 \to G/\Lambda$ and $\mathcal{S}_2 \to G/\Lambda$ are injective and surjective, respectively. Then, there exists a set $\mathcal{F}$ in $B(G)$ which is a fundamental domain for the action of Λ on G satisfying $\mathcal{S}_1 \subset \mathcal{F} \subset \mathcal{S}_2$.*

We will construct $\mathcal{S}_1$ and $\mathcal{S}_2$ in the following subsections.

Remark 4.4. The constructed $\mathcal{S}_1$ and $\mathcal{S}_2$ will not strictly be Siegel sets of the form $\omega T_c K$, but rather of the form $\omega T_c K'$ for some subset K' of K. We will call them Siegel sets regardless.

4.3.1. *Constructing $\mathcal{S}_1$.* To obtain the domain $\mathcal{S}_1$, we will use general properties of the Borel–Serre compactification following [10]. The construction below holds for a general connected semisimple algebraic group G over $\mathbb{Q}$, unless otherwise specified (note that our group G is always semisimple by [21, Proposition 3.7]).

Consider the symmetric space $X = G(\mathbb{R})/K$, where K is a maximal compact subgroup of $G(\mathbb{R})$. For each parabolic $\mathbb{Q}$-subgroup P of G, let $S_P := (R_dP/(R_uP \cdot R_dG))$, where R_u denotes the unipotent radical and R_d denotes the $\mathbb{Q}$-split part. Then, S_P is a $\mathbb{Q}$-split torus, and we let $A_P := S_P(\mathbb{R})^\circ$. There is a natural action of A_P on X called the geodesic action (see [10, (3.2)]). Set $e(P) = A_P \backslash X$, and consider

$$\overline{X} = \coprod_{P \text{ parabolic}} e(P),$$

which by [10, (7.1)] naturally has a structure of a manifold with corners. The topology of $\overline{X}$ is studied in [10, Sections 5, 6]; in particular, it is shown that for any parabolic group P, the subset $X(P) = \coprod_{Q \supset P} e(Q)$ is an open subset of $\overline{X}$. Taking $P = G$, we see that $e(G) = X$ is an open submanifold of $\overline{X}$.

Assume for simplicity that G is split over $\mathbb{Q}$ with split maximal torus T. Let $P = \overline{N}T$ be a Borel subgroup of G. For $x \in X$ and a real constant $c > 0$, we can consider the set

$$U_{x,P,c} = \overline{N}(\mathbb{R})(T_c \cdot x).$$

Its closure $\overline{U_{x,P,c}}$ in $\overline{X}$ is a neighbourhood of the closure of $e(P)$ in $\overline{X}$. Then, we have the following result (see [10, Proposition 10.3]):

Proposition 4.5. *There exists $c > 0$ satisfying that for any $g_1, g_2 \in \overline{U_{x,P,c}}$, if there exists $\gamma \in G(\mathbb{Z})$ such that $g_1 = \gamma g_2$, then $\gamma \in P(\mathbb{Z})$.*

To obtain a suitable Siegel set $\mathcal{S}_1$, we need to carefully choose a compact subset $\omega \subset \overline{N}(\mathbb{R})$. Let $(\alpha_1, \dots, \alpha_k)$ be an ordering of the positive roots of G satisfying that $\mathrm{ht}(\alpha_i) \leq \mathrm{ht}(\alpha_{i+1})$ for all $1 \leq i \leq k-1$. For each root α_i we consider the isomorphism $u_{\alpha_i} \colon \mathbb{G}_a \to U_{\alpha_i}$, where $U_{\alpha_i} \subset \overline{N}$. By [12, Theorem 5.1.13], there is an isomorphism of varieties over $\mathbb{Z}$:

$$\prod_{i=1}^{k} U_{\alpha_i} \longrightarrow \overline{N}$$

which is just the multiplication map. In other words, we can express any element of $\overline{N}(\mathbb{R})$ as $u_{\alpha_1}(x_1) \cdots u_{\alpha_k}(x_k)$ for some $x_1, \dots, x_k \in \mathbb{R}$. Moreover, a set of $x_1, \dots, x_k$ will correspond to an element of $\overline{N}(\mathbb{Z})$ if and only if $x_1, \dots, x_k \in \mathbb{Z}$. We now recall the following result (see e.g. [12, Proposition 5.1.14]):

Lemma 4.6. *Let $x, y \in \mathbb{R}$, and let α, β be positive roots. Then,*

$$u_\alpha(x) u_\beta(y) u_\alpha(-x) u_\beta(-y) = \prod_{i,j>0} u_{i\alpha+j\beta}(c_{i,\alpha,j,\beta} x^i y^j).$$

Here $c_{i,\alpha,j,\beta}$ is a constant, and the product is taken over all $i, j > 0$ such that $i\alpha + j\beta$ is a positive root.

Consider the set $\varpi = \{u_{\alpha_1}(x_1) \cdots u_{\alpha_k}(x_k) \in \overline{N}(\mathbb{R}) \mid x_i \in [-1/2, 1/2] \ \forall\, i\}$ inside $\overline{N}(\mathbb{R})$.

Proposition 4.7. *We have that*

1. $\overline{N}(\mathbb{Z})\varpi = \overline{N}(\mathbb{R})$.
2. *Except for a set of zero measure, no two distinct elements of ϖ are $\overline{N}(\mathbb{Z})$-translates of each other.*

Proof. For the first point, let $y_1, \dots, y_k \in \mathbb{R}$. We will show that there exist $n_1, \dots, n_k \in \mathbb{Z}$ and $x_1, \dots, x_k \in \mathbb{R}$ such that

$$u_{\alpha_1}(n_1) \cdots u_{\alpha_k}(n_k) u_{\alpha_1}(x_1) \cdots u_{\alpha_k}(x_k) = u_{\alpha_1}(y_1) \cdots u_{\alpha_k}(y_k). \tag{4.1}$$

Using the commutator relations of Lemma 4.6, we can reorder the terms in the left hand side to get equations of the form

$$y_m = n_m + x_m + p_m(n_1, \dots, n_k, x_1, \dots, x_k), \tag{4.2}$$

where p_m are polynomials. By examining the commutator relations, we see that p_m only depends on the variables corresponding to lower height coefficients. In particular, if α_m is a height-one root, we can choose $n_m \in \mathbb{Z}$

and $x_m \in [-1/2, 1/2]$ such that $y_m = n_m + x_m$. We can then find coefficients n_m, x_m for the larger height roots inductively using (4.2).

For the second point, choose two elements of $\overline{\omega}$ with coefficients $x_1, \dots, x_k$ and $y_1, \dots, y_k$ lying in $(-1/2, 1/2)$. Assume there exist $n_1, \dots, n_k \in \mathbb{Z}$ satisfying (4.1). By induction, we will show that $n_i = 0$ for all i. This is clear for the height-one coefficients, since $n_i + x_i = y_i$. Assume by induction that all the coefficients n_i are zero up to some height h. We note that by Lemma 4.6 all terms in the polynomial $p_m(n_1, \dots, n_k, x_1, \dots, x_k)$ are multiple of at least one n_i of lower height. Hence, by induction we get that $p_m(n_1, \dots, n_k, x_1, \dots, x_k) = 0$ and thus that $n_m = 0$, as wanted. □

Assume from now on that our group G is one of the groups constructed in Section 2.1. We note that $T^\theta(\mathbb{Z})$ acts by conjugation on $\overline{\omega}$: recall that $\mathrm{Ad}(t) \cdot u_\alpha(x) = u_\alpha(\alpha(t) \cdot x)$, and for any $t \in T^\theta(\mathbb{Z})$ we have that $\alpha(t) = \pm 1$. Alternatively, we can say that there is a mapping $T^\theta(\mathbb{Z}) \to \{\pm 1\}^{\#S_G}$ given by $t \mapsto (\alpha(t))_{\alpha \in S_G}$; however, it needs not be surjective: denote by $A = \{a_1, \dots, a_l\}$ a set of representatives of the cokernel of this map. For any element $a_i \in A$, write it as $a_i = (a_{i,1}, \dots, a_{i,k})$, where $a_{i,j} = \pm 1$ correspondingly. Consider the set ω_i inside $\overline{\omega}$ consisting of those elements $u = u_{\alpha_1}(x_1) \cdots u_{\alpha_k}(x_k)$ such that for all height-one coefficients α_j, we have $x_j \in a_{i,j} \cdot [0, 1/2]$. Finally, define ω to be the union of the sets ω_i. Then, each element in $\overline{\omega}$ is conjugate to a unique element in ω.

Additionally, we note that $T^\theta(\mathbb{Z}) \subset K$, since $T^\theta(\mathbb{Z})$ is fixed by the Cartan involution τ chosen in Section 4.2, and $\tau|_{T^\theta}$ is just the inverse map.

Take $\mathcal{S}_1 = \omega T_c \overline{K}$, where $c > 0$ satisfies the conclusions of Proposition 4.5, and $\overline{K}$ is a fundamental set for the action of $Z(G)(\mathbb{Z})$ on K. Let $g_1 = n_1 t_1 k_1$ and $g_2 = n_2 t_2 k_2$ be two elements of $\mathcal{S}_1$, and moreover we assume that n_1 and n_2 lie in the interior of ω (this interior is a set of full measure). Assume that g_1 and g_2 are equivalent under the $G(\mathbb{Z})$-action. By Proposition 4.5, it follows that g_1 and g_2 have to be $P^\theta(\mathbb{Z})$-conjugate, say by an element $p_0 = n_0 t_0$ for $n_0 \in \overline{N}(\mathbb{Z})$ and $t_0 \in T^\theta(\mathbb{Z})$. Then, we can write

$$n_0(t_0 n_1 t_0^{-1}) t_1 (t_0 k_1) = n_2 t_2 k_2.$$

By uniqueness in the Iwasawa decomposition, we have that $n_0(t_0 n_1 t_0^{-1}) = n_2$, $t_1 = t_2$ and $t_0 k_1 = k_2$. If we look at the first equation in terms of height-one roots α_i, we get equalities of the form $u_{\alpha_i}(x_0) u_{\alpha_i}(\alpha_i(t_0) x_1) = u_{\alpha_i}(x_2)$, where $x_0 \in \mathbb{Z}$ and $x_1, x_2 \in [-1/2, 1/2]$ (or a subinterval if appropriate). This can only happen if $x_0 = 0$ for all coefficients, meaning that $n_0 = 1$, and also by construction of ω it must also happen that $\alpha_i(t_0) = 1$ for all i, or in other words that $t_0 \in Z(G)(\mathbb{Z})$. Then, the last equation $t_0 k_1 = k_2$ can only happen if $t_0 = 1$ by construction. Therefore, $g_1 = g_2$ as wanted.

4.3.2. ***Constructing $\mathcal{S}_2$.*** We can construct $\mathcal{S}_2$ compatibly with $\mathcal{S}_1$ thanks to the following proposition:

Proposition 4.8. *There exists a real constant $c > 0$ such that $G(\mathbb{R}) = G(\mathbb{Z})\omega T_c \overline{K}$, where ω and $\overline{K}$ are as in Section 4.3.1.*

Proof. We can show that $G(\mathbb{R}) = G(\mathbb{Z})\omega' T_c K$ for some compact subset $\omega' \subset N(\mathbb{R})$ and some $c > 0$ using [25, Theorem 4.15], the first statement is reduced to showing that $G(\mathbb{Q}) = P^\theta(\mathbb{Q})G(\mathbb{Z})$, which follows from [9, Section 6, Lemma 1(b)].

It is clear that K can be substituted by $\overline{K}$, since we can multiply by an appropriate element of $Z(G)(\mathbb{Z})$ in $G(\mathbb{Z})$. Now, let $g = g_0 n t k$ be an element of $G(\mathbb{R}) = G(\mathbb{Z})\omega' T_c \overline{K}$: we will show that $g \in G(\mathbb{Z})\omega T_c \overline{K}$. We know that there exists $n_0 \in N(\mathbb{Z})$ and $t_0 \in T^\theta(\mathbb{Z})$ such that $t_0 n_0 n t_0^{-1} \in \omega$. Let $z \in Z(G)(\mathbb{Z})$ be such that $z t_0 k \in \overline{K}$. Then, we can write $g = (z^{-1} g_0 n_0^{-1} t_0^{-1})(t_0 n_0 n t_0^{-1}) t (z t_0 k) \in G(\mathbb{Z})\omega T_c \overline{K}$, as wanted. □

We fix $\mathcal{S}_2 = \omega T_c \overline{K}$, for some $c > 0$ satisfying the above proposition. It is clear then that $\mathcal{S}_1$ and $\mathcal{S}_2$ satisfy the required properties, and hence that by Lemma 4.3 we obtain a box-shaped fundamental domain $\mathcal{F}$ for the action of $G(\mathbb{Z})$ over $G(\mathbb{R})$.

5. Counting reducible orbits

In light of the results of Section 3, to estimate the elements of $B(\mathbb{Z})$ having discriminant divisible by the square of a large prime, it suffices to count certain special reducible $G(\mathbb{Z})$-orbits in $V(\mathbb{Z})$. In this section, we develop much of what we will need in this regard, following Bhargava's geometry-of-numbers techniques, and in particular using the ideas in [31].

5.1. Averaging. Let $S \subset V(\mathbb{Z})$ be a $G(\mathbb{Z})$-invariant subset. Define

$$N(S, X) = \sum_{\substack{v \in G(\mathbb{Z})\backslash S \\ \mathrm{ht}(v) < X}} \frac{1}{\#\,\mathrm{Stab}_G(v)(\mathbb{Z})}.$$

We will prove the following:

Theorem 5.1. *There exist real positive constants C, δ such that*

$$N(V(\mathbb{Z})^{red}, X) = C X^{\dim V} + O(X^{\dim V - \delta}).$$

By analogous arguments to [38, Section 2.9], there exist open subsets $L_1, \ldots, L_k$ covering $\{b \in B(\mathbb{R}) \mid \mathrm{ht}(b) = 1, \Delta(b) \neq 0\}$ such that for a fixed i, the quantity $r_i = \#\,\mathrm{Stab}_{G(\mathbb{R})}(v)$ remains constant for any choice of $v \in \pi^{-1}(L_i)$. We will denote $\Lambda = \mathbb{R}_{>0}$ and $V_i := V(\mathbb{Z})^{red} \cap G(\mathbb{R})\kappa(\Lambda L_i)$. Fix a compact left and right K-invariant set $G_0 \subset G(\mathbb{R})$ which is the closure of a non-empty open set, for which we assume that $G_0 = G_0^{-1}$. An averaging argument just as in [6, Section 2.3] yields

$$N(V_i, X) = \frac{1}{r_i \,\mathrm{vol}(G_0)} \int_{g \in \mathcal{F}} \#\{v \in V(\mathbb{Z})^{red} \cap (g G_0 \kappa(\Lambda L_i))_{<X}\} \mathrm{d}g. \tag{5.1}$$

To obtain the estimate for $N(V(\mathbb{Z})^{red}, X)$, it will suffice to obtain the appropriate estimates for $N(V_i, X)$. For any subset S inside $V(\mathbb{Z})^{red} \cap G(\mathbb{R})\kappa(\Lambda L_i)$, we can use the expression (5.1) to *define* $N(S, X)$ as

$$N(S,X) = \frac{1}{r_i \operatorname{vol}(G_0)} \int_{g\in\mathcal{F}} \#\{v \in S \cap (gG_0\kappa(\Lambda L_i))_{<X}\}\mathrm{d}g.$$

For the argument, it will be crucial to use Davenport's lemma (see [13]), as stated in [6, Proposition 2.6]. We record it here for convenience.

Proposition 5.2. *Let $\mathcal{R}$ be a bounded, semialgebraic multiset in $\mathbb{R}^n$ having maximum multiplicity m and that is defined by at most k polynomial inequalities, each having degree at most l. Then,*

$$\#(\mathcal{R} \cap \mathbb{Z}^n) = \operatorname{vol}(\mathcal{R}) + O(\max(\{\operatorname{vol}(\overline{\mathcal{R}}), 1\})),$$

where $\operatorname{vol}(\overline{\mathcal{R}})$ *denotes the greatest d-dimensional volume of any projection of $\mathcal{R}$ onto a coordinate subspace obtained by equating $n-d$ coordinates to zero, and where d takes any value between 1 and $n-1$. The implied constant in the second summand depends only on n, m, k and l.*

5.2. Applying the Selberg sieve. Another important step in our argument will be the use of the Selberg sieve. Notably, in the statement of Theorem 5.1 we require a power saving estimate in the error term, which we will obtain by applying the Selberg sieve as in [32]. In this section, we describe exactly how the Selberg sieve is used, and which hypothesis are needed.

The general situation is the following: suppose we have a finite sequence of non-negative numbers $\mathcal{A} = (a_n)_n$, and let P be a finite product of distinct primes. For all $d|P$, assume the following holds:

$$\sum_{n\equiv 0 \bmod d} a_n = g(d)X + r_d, \tag{5.2}$$

where $X > 0$ and $g(d)$ is a multiplicative function satisfying $0 < g(p) < 1$ for all primes $p|P$. Define the multiplicative function h by $h(p) = \frac{g(p)}{1-g(p)}$ at primes p. For some choice of $D_0 > 1$, write

$$H = \sum_{\substack{d<\sqrt{D_0}\\ d|P}} h(d).$$

Then, [17, Theorem 6.4] says that

$$\sum_{(n,P)=1} a_n \le XH^{-1} + O\left(\sum_{\substack{d\le D_0\\ d|P}} \tau_3(d) r_d\right). \tag{5.3}$$

We now explain how to apply (5.3) in our context of orbit-counting. We will typically work in a subset $W \subset V(\mathbb{Z})$ (e.g. the main body, the cusp...), and we will suppose we have a set $S \subset W$ which satisfies $S = \cap_p S_p$, where for each prime p, the set S_p is defined by congruence conditions modulo p. We wish to estimate $N(S, X)$, which will generally be some orbit-counting function of S inside W (to be made precise in future applications). Let T_p be the complement of S_p in W, and fix a number $z < X$. Let $P(z) = \prod_{p<z} p$, and for a number $d|P(z)$ set

$$a_d = N\left(\bigcap_{p|d} T_p \bigcap_{p|\frac{P(z)}{d}} S_p, X\right).$$

If $d \nmid P(z)$, set $a_d = 0$. To apply the Selberg sieve, we need an estimate like (5.2). Let L be a translate of mW, for some squarefree $m \in \mathbb{Z}$, and assume that we have an estimate of the form

$$N(L, X) = km^{-A}X^B + O(m^{-A+C}X^{B-D}), \tag{5.4}$$

for some non-negative constants A, B, C, D and k. Then, it follows that

$$\sum_{n \equiv 0 \bmod d} a_n = N(\cap_{p|d} T_p, X) = kg_d X^B + r_d,$$

where for a prime p, the quantity g_p is the density of T_p, for d squarefree we set $g_d = \prod_{p|d} g_p$, and we have $r_d = O(d^C g_d X^{B-D})$. Then, by (5.3), we have

$$a_1 = \sum_{(n,P(z))=1} a_n \leq kX^B H^{-1} + O\left(\sum_{\substack{d \leq D_0 \\ d|P}} \tau_3(d) r_d\right).$$

Assume now that as $p \to \infty$, the density g_p converges to some constant $\lambda \in (0, 1)$. Then, we are able to obtain bounds for H and r_d depending only on X and the choice of D. Given that $d^{-\varepsilon} \ll_\varepsilon g_d \ll_\varepsilon d^\varepsilon$, we get that $H = D_0^{1/2+o(1)}$. For the error term, we get that

$$\left|\sum_{\substack{d \leq D_0 \\ d|P}} \tau_3(d) r_d\right| \ll_\varepsilon X^{B-D} D_0^\varepsilon \sum_{d \leq D_0} d^C \ll_\varepsilon X^{B-D} D_0^{C+1+\varepsilon}.$$

The end result is that $a_1 \ll_\varepsilon X^B D_0^{-1/2+\varepsilon} + X^{B-D} D_0^{C+1+\varepsilon}$. By making an appropriate choice of D as a power of X, we can optimise this expression to yield $a_1 = O(X^{B-\delta})$ for some $\delta > 0$.

So, in summary, to use the Selberg sieve in the same way that is used in [32], it will suffice to have an expression of the form (5.4), and a proof

that the densities of our sets S_p converge to some constant in $(0,1)$ as p goes to infinity.

5.3. Reductions. We return to the setting of Section 5.1, where we had

$$N(V_i, X) = \frac{1}{r_i \operatorname{vol}(G_0)} \int_{g \in \mathcal{F}} \#\{v \in V(\mathbb{Z})^{red} \cap (gG_0\kappa(\Lambda L_i))_{<X}\} \mathrm{d}g.$$

To estimate this quantity, we will make some necessary reductions. We will begin with a "cutting-off-the-cusp" result, which amounts to saying that not too many points in the cusp are irreducible.

Proposition 5.3. *Let v_0 be the coefficient of the highest weight in V. Then, there exists a constant $\delta_1 > 0$ such that*

$$\int_{g \in \mathcal{F}} \#\{v \in (V(\mathbb{Z}) \setminus W_0(\mathbb{Z})) \cap g\mathcal{B}_X \mid v_0 = 0\} \mathrm{d}g = O(X^{\dim V - \delta_1}).$$

Proof. This is the content of [21, Proposition 8.12]. □

In a similar spirit, we also show that most of the elements in the main body are irreducible:

Proposition 5.4. *Let v_0 be the coefficient of the highest weight in V. Then, there exists a constant $\delta_2 > 0$ such that*

$$\int_{g \in \mathcal{F}} \#\{v \in V(\mathbb{Z})^{red} \cap g\mathcal{B}_X \mid v_0 \neq 0\} dg = O(X^{\dim V - \delta_2})$$

Proof. We will prove this statement by using the Selberg sieve, as explained in Section 5.2. First of all, if $v \in V(\mathbb{Z})$ is reducible, then for all primes p not dividing the N fixed in Proposition 2.13 the reduction of v mod p is reducible, since by Theorem 2.14 v is $G(\mathbb{Z}_p)$ conjugate to κ_b. By [21, Proof of Lemma 8.22], the density of elements in $V(\mathbb{F}_p)$ which are $\mathbb{F}_p$-reducible converges to some constant $\lambda \in (0,1)$.

To apply the Selberg sieve, we need some result in the style of (5.4). This is essentially the content of [21, Proposition 8.15 and Theorem 8.17]; a power saving estimate can be obtained similarly to [4, Proposition 10.5], and the contribution from the congruence conditions can be done similarly to our proof of Theorem 6.1: we do not repeat it here for the sake of concision. □

5.4. Counting reducible orbits. The previous reductions show that when trying to estimate

$$N(V_i, X) = \frac{1}{r_i \operatorname{vol}(G_0)} \int_{g \in \mathcal{F}} \#\{v \in V(\mathbb{Z})^{red} \cap (gG_0\kappa(\Lambda L_i))_{<X}\} \mathrm{d}g,$$

it is sufficient to work over the cusp $W_0(\mathbb{Z})$ up to a power-saving error term. Given that $\mathcal{F}$ is a box-shaped fundamental domain, we can write it as a

disjoint union $\mathcal{F}' \cup \omega T_c \overline{K}$, where ω, c and $\overline{K}$ are as in Section 4.3.1 and $\mathcal{F}'$ is a subset of

$$\omega \cdot \{t \in T^\theta(\mathbb{R})^\circ \mid \alpha(t) \leq c \text{ for some } \alpha \in S_G\} \cdot \overline{K}.$$

An explicit computation (e.g. following the reasoning in this section and in Section 5.6) shows that the integral in (5.1) is negligible when $\mathcal{F}$ is substituted by $\mathcal{F}'$. Hence, it suffices to integrate over $\omega T_c \overline{K}$. Given that G_0 is K-invariant and that dk can be normalised so that $\overline{K}$ has volume 1, we get:

$$N(V_i, X) = \frac{1}{r_i \operatorname{vol}(G_0)} \int_{n \in \omega} \int_{t \in T_c} \#\{v \in W_0(\mathbb{Z}) \cap nt\mathcal{B}_X\} \delta^{-1}(t) \mathrm{d}n \mathrm{d}^\times t + O\left(X^{\dim V - \delta}\right)$$

for some $\delta > 0$. It would be desirable to estimate the lattice points in the region using Davenport's lemma; however, as noted in [31], the cuspidal region is too skewed to apply the lemma directly: in particular, some of volumes of the projections can be of the order of the main term. To circumvent this, we will "slice" the region $W_0(\mathbb{Z})$ according to the values of the height-one coefficients. For $v \in W_0(\mathbb{Z})$, denote by $(\sigma_1(v), \dots, \sigma_r(v))$ its height-one coefficients. Then, for any $b = (b_1, \dots, b_r) \in \mathbb{R}^r$ and any subset $S \subset W_0(\mathbb{R})$, we will denote

$$S|_b = \{v \in S \mid (\sigma_1(v), \dots, \sigma_r(v)) = b\}.$$

Then, we can express

$$\#(W_0(\mathbb{Z}) \cap (nt\mathcal{B}_X)) = \sum_{b \in \mathbb{Z}^r} \#(W_0(\mathbb{Z}) \cap (nt\mathcal{B}_X)|_b).$$

Actually, we can assume that in the sum over $b = (b_1, \dots, b_r) \in \mathbb{Z}^r$, none of the components b_i are equal to zero due to the following:

Lemma 5.5. *Let $v \in W_0(\mathbb{R})$. If $\sigma_i(v) = 0$ for some i, then $\Delta(v) = 0$.*

Proof. Let $\{\alpha_1, \dots, \alpha_k\}$ be the height-one weights, and assume that the coefficient of α_i of v is zero. Let $\lambda_i \colon \mathbb{G}_m \to G_{\mathbb{C}}$ be the one-parameter subgroup such that $(\alpha_j \circ \lambda_i)(t) = t^{\delta_{ij}}$. Then, v has no positive weights with respect to λ_i, and so by Proposition 2.2 we get the result. □

When applying Proposition 5.2 to $(nt\mathcal{B}_X)|_b$, we get

$$\#(W_0(\mathbb{Z}) \cap (nt\mathcal{B}_X)|_b) = \operatorname{vol}((nt\mathcal{B}_X)|_b)(1 + O(X^{-1})). \tag{5.5}$$

The term $O(X^{-1})$ can be obtained as follows: each coefficient v_0 in $W_0(\mathbb{R})$ has a weight under the action of T, which we will denote $w(v_0)$. When performing the slicing, that is, fixing the values of the height-one coefficients,

all the weights turn out to be $\gg X$, and given that the volume of the region is the product of the weights of the different coordinates, we obtain the saving of size X.

Given that unipotent transformations preserve the volume, and that we can normalise $\mathrm{d}n$ so that $\mathrm{vol}(\omega) = 1$, we can write the following:

$$N(V_i, X) = \frac{1}{r_i \,\mathrm{vol}(G_0)} \sum_{b \in (\mathbb{Z}\setminus\{0\})^r} \int_{t \in T_c} \mathrm{vol}((t\mathcal{B}_X)|_b)\delta^{-1}(t)\mathrm{d}^\times t + O(X^{\dim V - \delta}). \tag{5.6}$$

For each height-one coefficient v_i, we will denote $\beta_i := (Xw(v_i)b_i)^{-1}$, and $\beta = (\beta_i)_i$. Denote by $W_\flat$ the set of coordinates of W_0 of non-positive height. It follows that

$$\mathrm{vol}((t\mathcal{B}_X)|_b) = \mathrm{vol}(tX \cdot \mathcal{B}|_\beta) = X^{\dim W_\flat} \prod_{v \in W_\flat} w(v)\,\mathrm{vol}(\mathcal{B}|_\beta).$$

We will make the change of variables $t \mapsto \beta = (\beta_1, \dots, \beta_r)$, under which $\mathrm{d}^\times t = \mathrm{d}^\times \beta = \prod_i \frac{\mathrm{d}\beta_i}{\beta_i}$. In Section 5.6, we will explicitly compute the volume of the cuspidal region for each of the possible cases. We will obtain a polynomial $Z(\beta) = \prod_i \beta_i^{e_i}$ with integer exponents $e_i \geq 2$, and we will see that

$$X^{\dim W_\flat} \prod_{v \in W_\flat} w(v)\delta^{-1}(t) = X^{\dim V} \frac{Z(\beta)}{Z(b)}. \tag{5.7}$$

It follows that

$$\int_{t \in T_c} \mathrm{vol}((t\mathcal{B}_X)|_b)\delta^{-1}(t)\mathrm{d}^\times t = \frac{X^{\dim V}}{|Z(b)|} \int_{\beta \in \mathbb{R}^r_{>0} \setminus T'} Z(\beta)\,\mathrm{vol}(\mathcal{B}|_\beta)\mathrm{d}^\times \beta,$$

where T' is the region corresponding to $T(\mathbb{R}) \setminus T_c$. It is not difficult to see that the integral over T' is $O(X^{-1})$, and hence can be added to the error term. For an element $v \in W_0(\mathbb{Z})$, define $Z(v) := Z(\sigma_i(v)) = \prod_i \sigma_i(v)^{e_i}$ and $Z^\times(v) := \prod_i \sigma_i(v)^{e_i - 1}$. Then,

$$\begin{aligned} \int_{t \in T_c} \mathrm{vol}((t\mathcal{B}_X)|_b)\delta^{-1}(t)\mathrm{d}^\times t &= \frac{X^{\dim V}}{|Z(b)|} \int_{\beta \in \mathbb{R}^r_{>0}} Z(\beta)\,\mathrm{vol}(\mathcal{B}|_\beta)\mathrm{d}^\times \beta + O(X^{\dim V - 1}) \\ &= \frac{X^{\dim V}}{|Z(b)|} \int_{v \in \mathcal{B} \cap W_0(\mathbb{R})_+} Z^\times(v)\mathrm{d}v + O(X^{\dim V - 1}). \end{aligned} \tag{5.8}$$

Here, $W_0(\mathbb{R})_+ = \{v \in W_0(\mathbb{R}) \mid \sigma_i(v) > 0, \forall\, i\}$. Combining (5.6) and (5.8), and summing over all b, we obtain that

$$N(V_i, X) = 2^r \prod_{i=1}^{r} \zeta(e_i) \cdot \left(\frac{1}{r_i \operatorname{vol}(G_0)} \int_{v \in \mathcal{B} \cap W_0(\mathbb{R})_+} Z^{\times}(v) \mathrm{d}v \right) X^{\dim V} + O(X^{\dim X - \delta}). \tag{5.9}$$

To obtain the desired asymptotic for $N(V(\mathbb{Z})^{red}, X)$, it suffices to use the inclusion-exclusion principle. For any subset $I \subset \{1, \dots, r\}$, the same procedure as above obtains (5.9) for the set $V_I = \bigcap_{i \in I} V_i$, with the appropriate constants substituted. This concludes the proof of Theorem 5.1.

5.5. Computing the constant. As promised, we will compute the constant of the main term of Theorem 5.1. We will do so using a Jacobian change-of-variables formula, whose statement and proof are completely analogous to [31, Proposition 14]: we include the proof in our case for convenience.

Proposition 5.6. *Let $\phi\colon W_0(\mathbb{R}) \to \mathbb{R}$ be a measurable function. Then, there exists a non-zero rational constant $\mathcal{J} \in \mathbb{Q}^{\times}$ such that*

$$\int_{v \in W_0(\mathbb{R})} \phi(v) |Z^{\times}(v)| \mathrm{d}v = |\mathcal{J}| \int_{\substack{b \in B(\mathbb{R}) \\ \Delta(b) \neq 0}} \left(\sum_{v \in \frac{\pi^{-1}(b)}{P^{-,\theta}(\mathbb{R})}} \int_{h \in P^{-,\theta}(\mathbb{R})} \phi(h \cdot v) \mathrm{d}h \right) \mathrm{d}b.$$

Here, $\mathrm{d}v$ and $\mathrm{d}b$ are Euclidean measures, and $dh = \delta^{-1}(t) \mathrm{d}n \mathrm{d}^{\times}t$ is a right Haar measure for $P^{-,\theta}(\mathbb{R})$.

Proof. Let $U \subset B(\mathbb{R})$ be an open subset, and let $\sigma\colon U \to W_0(\mathbb{R})$ be a continuous section of the GIT quotient map $\pi\colon V \to B$. We first claim that we have

$$\int_{v \in P^{-,\theta}(\mathbb{R})\sigma(U)} \phi(v) |Z^{\times}(v)| \mathrm{d}v = |\mathcal{J}| \int_{b \in U} \int_{h \in P^{-,\theta}(\mathbb{R})} \phi(h \cdot \sigma(b)) \mathrm{d}h \mathrm{d}b \tag{5.10}$$

for some non-zero rational constant $\mathcal{J}$. By the Stone–Weierstrass theorem, we can assume that σ is piecewise analytic, in which case we have

$$\int_{v \in P^{-,\theta}(\mathbb{R})\sigma(U)} \phi(v) |Z^{\times}(v)| \mathrm{d}v = \int_{b \in U} \int_{h \in P^{-,\theta}(\mathbb{R})} |\mathcal{J}_{\sigma}(h, b)| \phi(h \cdot \sigma(b)) \mathrm{d}h \mathrm{d}b,$$

where $\mathcal{J}_{\sigma}(h, b)$ denotes the determinant of the Jacobian matrix arising from the change of variables that takes the measure $Z^{\times}(v)\mathrm{d}v$ to $\mathrm{d}h\mathrm{d}b$. We will now show that $\mathcal{J}_{\sigma}(h, b)$ is independent of σ, h and b.

To show that $\mathcal{J}_{\sigma}(h, b)$ is independent of h, we fix $\gamma \in P^{-,\theta}(\mathbb{R})$ and consider the change of variables $v \mapsto \gamma \cdot v$ in $W_0(\mathbb{R})$. We have that $Z^{\times}(\gamma \cdot v)\mathrm{d}(\gamma \cdot v) = \chi(\gamma) Z^{\times}(v)\mathrm{d}v$ for some character $\chi\colon P^{-,\theta}(\mathbb{R}) \to \mathbb{R}_{>0}$,

which we now determine explicitly. If $\gamma \in \overline{N}(\mathbb{R})$, then $\chi(\gamma) = 1$, since neither $Z^\times$ or the volume of $W_0(\mathbb{R})$ are changed by the action of $\overline{N}(\mathbb{R})$. Now, assume that $\gamma \in T^\theta(\mathbb{R})$. On one hand, we have that

$$Z^\times(\gamma \cdot v) = \left(\prod_{\alpha_i \in S_G} \alpha_i(\gamma)^{e_i - 1} \right) Z^\times(v).$$

On the other hand, we have that

$$d(\gamma \cdot v) = \left(\prod_{\substack{\alpha \in \Phi(G,T^\theta) \\ \mathrm{ht}(\alpha) \leq 1}} \alpha(\gamma) \right)^{-1} \mathrm{d}v.$$

In view of (5.7), we conclude that $\chi(\gamma) = \delta^{-1}(\gamma)$. On the other hand, for $\gamma = nt \in P^\theta(\mathbb{R})$ we also have that

$$\mathcal{J}_\sigma(\gamma h, b) d(\gamma h) \mathrm{d}b = \delta^{-1}(t) \mathcal{J}_\sigma(\gamma h, b) \mathrm{d}h \mathrm{d}b$$

because $\mathrm{d}h$ is a right Haar measure of H, and $\delta^{-1}(t)$ is the corresponding modular function (cf. [19, (8.26)]). We then have that

$$\mathcal{J}_\sigma(\gamma h, b) d(\gamma h) db = Z^\times(\gamma v) \mathrm{d}v = \delta^{-1}(t) Z^\times(v) \mathrm{d}v = \delta^{-1}(t) \mathcal{J}_\sigma(h, b) \mathrm{d}h \mathrm{d}b,$$

and hence that $\mathcal{J}_\sigma(h, b) = \mathcal{J}_\sigma(\gamma h, b)$ is independent of h, as wanted.

The rest of the proof now follows analogously to [6, Proof of Proposition 3.10]. More precisely, that $\mathcal{J}_\sigma(h, b)$ is independent of σ is analogous to Step 2 in [6, Proof of Proposition 3.10]; in particular, we can take σ to be the Kostant section. Then, independence of b follows from steps 3 and 4 in [6, Proof of Proposition 3.10].

Thus, we have shown (5.10). The proposition now follows from (5.10) in a similar way as how [6, Proposition 3.7] follows from [6, Proposition 3.10]. □

The proof of Theorem 5.1 shows that the leading constant in the asymptotic for $N(V_i, X)$ is:

$$2^r \prod_{i=1}^{r} \zeta(e_i) \sum_{i=1}^{k} \frac{1}{r_i \operatorname{vol}(G_0)} \int_{v \in \mathcal{B} \cap W_0(\mathbb{R})_+} Z^\times(v) \mathrm{d}v.$$

Here, r is the amount of height-one coefficients, ζ is the Riemann zeta function, and e_i are the exponents corresponding to $Z(\beta) = \prod_i \beta_i^{e_i}$. We will now give a more succinct description of the above integral, following [31, Section 4.3].

Given that G_0 is K-invariant, we can write it as $G_0 = \mathcal{S} \cdot K$, for some set $\mathcal{S} \subset P^{-,\theta}(\mathbb{R})$. We have the following lemma:

Lemma 5.7. *The map* $\pi \colon K\kappa(\Lambda L_i) \cap W_0(\mathbb{R})_+ \to L_i$ *is* r_i *to* 1.

Proof. The result follows from the fact that every element $v \in K\kappa(\Lambda L_i)$ satisfies $\#\operatorname{Stab}_{G(\mathbb{R})} v = r_i$, and that if $g \in \operatorname{Stab}_{G(\mathbb{R})} v$, then writing $g = pk$ for $p \in P^{-,\theta}(\mathbb{R})$ and $k \in K$, we get that $kv = p^{-1}v$, so kv belongs to $K\kappa(\Lambda L_i) \cap W_0(\mathbb{R})_+$. Conversely, given that $P^{-,\theta}$ acts simply transitively on $W_0(\mathbb{R})$, any element in $K\kappa(L_i) \cap W_0(\mathbb{R})_+$ that is conjugate to v has to be of the form $kv = p'v$ for some $k \in K$ and $p' \in P^{-,\theta}$. □

Now, setting ϕ to be the indicator function of $\mathcal{B} \cap W_0(\mathbb{R})_+$ in Proposition 5.6, we obtain

$$\frac{1}{r_i \operatorname{vol}(G_0)} \int_{v \in \mathcal{B}_{<X} \cap W_0(\mathbb{R})_+} Z^{\times}(v)\mathrm{d}v = \frac{|\mathcal{J}| r_i \operatorname{vol}(K\mathcal{S}) \operatorname{vol}(\{b \in \Lambda L_i \mid \operatorname{ht}(b) < 1\})}{r_i \operatorname{vol}(\mathcal{S}K)}. \tag{5.11}$$

However, we observe that:

Lemma 5.8. *We have $\mathcal{S}K = K\mathcal{S}$.*

Proof. Recall that G_0 is left and right K-invariant and satisfies $G_0^{-1} = G_0$. Then,

$$K\mathcal{S} \subset K\mathcal{S}K = \mathcal{S}K = G_0 = G_0^{-1} = K\mathcal{S}^{-1}.$$

By uniqueness in the Iwasawa decomposition, we must have that $\mathcal{S} \subset \mathcal{S}^{-1}$, and symmetrically that $\mathcal{S} = \mathcal{S}^{-1}$. Therefore, $\mathcal{S}K = K\mathcal{S}^{-1} = K\mathcal{S}$, as wanted. □

We are left to deal with the volumes of the corresponding L_i terms, which we do using the inclusion-exclusion principle. The end result is

$$N(V(\mathbb{Z})^{red}, X) \sim 2^r \prod_{i=1}^{r} \zeta(e_i) |\mathcal{J}| \operatorname{vol}(\{b \in B(\mathbb{R}) \mid \operatorname{ht}(b) < 1\}) X^{\dim V}.$$

We can compare this result with the asymptotics for $N(V(\mathbb{Z})^{irred}, X)$, which can be read off [21, Theorem 8.8]. In there, one of the factors of the constant is related to the volume of $G(\mathbb{Z})\backslash G(\mathbb{R})$ with respect to a suitably normalised Haar measure, and can be done following [23] and [11], for instance. Surprisingly, we get that

$$\operatorname{vol}(G(\mathbb{Z})\backslash G(\mathbb{R})) = c \prod_{i=1}^{r} \zeta(e_i),$$

where c is the order of the fundamental group of $G_{\mathbb{C}}$ and e_i turn out to be the same exponents as above; in particular, the constants for the reducible and irreducible case appear to be the same up to some rational factor, thus answering Question 2 of [31] affirmatively for our representations (G, V).

However, the two methods of obtaining the constants appear to be fundamentally different, and we wonder if there is any "natural" explanation as to why they should give the same result.

5.6. Case-by-case analysis. In this section, we complete the proof of Theorem 5.1 by performing a case-by-case analysis. For the D_n and E_n cases, we will explicitly compute the dimension and volume of $W_\flat$ (which was defined to be the set of coefficients of W_0 of non-positive height), and the modular function $\delta(t) = \prod_{\beta\in\Phi_G^-} \beta(t) = \det \mathrm{Ad}(t)|_{\mathrm{Lie}\,\overline{N}(\mathbb{R})}$.

5.6.1. D_{2n+1}. The exposition in the D_n cases is inspired by [21, Appendix A] and [30, Section 7.2.1]. We start by describing explicitly the representation (G, V) of D_{2n+1} in the form given by Table 2.1.

Let $n \geq 2$ be an integer. Let U_1 be a $\mathbb{Q}$-vector space with a basis given by $\{e_1, \dots, e_n, u_1, e_n^*, \dots, e_1^*\}$, endowed with the symmetric bilinear form b_1 satisfying $b_1(e_i, e_j) = b_1(e_i, u_1) = b_1(e_i^*, e_j^*) = b_1(e_i^*, u_1) = 0$, $b_1(e_i, e_j^*) = \delta_{ij}$ and $b_1(u_1, u_1) = 1$ for all $1 \leq i, j \leq n$. In this case, given a linear map $A\colon U \to U$ we can define its *adjoint* as the unique map $A^*\colon U \to U$ satisfying $b_1(Av, w) = b_1(v, A^*w)$ for all $v, w \in U$. In terms of matrices, A^* corresponds to taking the reflection of A along its antidiagonal when working with the fixed basis. We can define $\mathrm{SO}(U_1, b_1) := \{g \in \mathrm{SL}(U_1) \mid gg^* = \mathrm{id}\}$, with a Lie algebra that can be identified with $\{A \in \mathrm{End}(U) \mid A = -A^*\}$.

Let U_2 be a $\mathbb{Q}$-vector space with basis $\{f_1, \dots, f_n, u_2, f_n^*, \dots, f_1^*\}$, with a similarly defined bilinear form b_2. Let $(U, b) = (U_1, b_1) \oplus (U_2, b_2)$. Let $H = \mathrm{SO}(U, b)$, and consider $\mathfrak{h} := \mathrm{Lie}\, H$. With respect to the basis

$$\{e_1, \dots, e_n, u_1, e_n^*, \dots, e_1^*, f_1, \dots, f_n, u_2, f_n^*, \dots, f_1^*\},$$

the adjoint of a block matrix according to the bilinear form b is given by

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^* = \begin{pmatrix} A^* & C^* \\ B^* & D^* \end{pmatrix},$$

where A^*, B^*, C^*, D^* denote reflection by the antidiagonal. An element of $\mathfrak{h}$ is given by

$$\left\{ \begin{pmatrix} B & A \\ -A^* & C \end{pmatrix} \,\middle|\, B = -B^*,\, C = -C^* \right\}.$$

The involution θ is given by conjugation with $\mathrm{diag}(1, \dots, 1, -1, \dots, -1)$, where the first $2n+1$ entries are 1 and the last $2n+1$ entries are given by -1. Under this description, we see that

$$V = \left\{ \begin{pmatrix} 0 & A \\ -A^* & 0 \end{pmatrix} \,\middle|\, A \in \mathrm{Mat}_{(2n+1)\times(2n+1)} \right\}.$$

Moreover, $G = (H^\theta)^\circ$ is isomorphic to $\mathrm{SO}(U_1) \times \mathrm{SO}(U_2)$. We will use the map

$$\begin{pmatrix} 0 & A \\ -A^* & 0 \end{pmatrix} \longmapsto A$$

to establish a bijection between V and $\mathrm{Hom}(U_2, U_1)$, where the action of $(g, h) \in \mathrm{SO}(U_1) \times \mathrm{SO}(U_2)$ on $A \in V$ is $(g, h) \cdot A = gAh^{-1}$.

Let T be the maximal torus of G given by

$$\mathrm{diag}(t_1, \ldots, t_n, 1, t_n^{-1}, \ldots, t_1^{-1}, s_1, \ldots, s_n, 1, s_n^{-1}, \ldots, s_1^{-1})$$

A basis of simple roots for G is

$$S_G = \{t_1 - t_2, \ldots, t_{n-1} - t_n\} \cup \{s_1 - s_2, \ldots, s_{n-1} - s_n\}.$$

A positive root basis for V can be taken to be

$$S_V = \{t_1 - s_1, s_1 - t_2, \ldots, t_n - s_n, s_n\}.$$

For convenience, we now switch to multiplicative notation for the roots. We make the change of variables $\alpha_i = t_i/t_{i+1}$ for $i = 1, \ldots, n-1$ and $\alpha_n = t_n$; similarly $\gamma_i = s_i/s_{i+1}$ for $i = 1, \ldots, n-1$ and $\gamma_n = s_n$. The estimate for the volume of $W_\flat$ becomes:

$$\prod_{v \in W_\flat} Xw(v) = X^{2n^2+2n+1} \prod_{i=1}^{n} \alpha_i^{-2in+i^2-2i} \gamma_i^{-2in+i^2}.$$

The modular function in our case is

$$\delta^{-1}(t) = \prod_{i=1}^{n} \alpha_i^{2in-i^2} \gamma_i^{2in-i^2}.$$

Changing variables to $\beta_i = (Xw(v_i)b_i)$, where v_i are the height-one coefficients, we obtain

$$\prod_{v \in W_\flat} Xw(v)\delta^{-1}(t) = X^{4n^2+4n+1} \frac{Z(\beta)}{Z(b)},$$

where $Z(\beta) := \prod_{i=1}^{n} (\beta_{2i-1}\beta_{2i})^{2i}$.

5.6.2. ***D_{2n}.*** The analysis in this case is very similar to the D_{2n+1} case. Now, we consider the $\mathbb{Q}$-vector space U_1 with basis $\{e_1, \ldots, e_n, e_n^*, \ldots, e_1^*\}$, endowed with a symmetric bilinear form satisfying $b_1(e_i, e_j) = b_1(e_i^*, e_j^*) = 0$, $b_1(e_i, e_j^*) = \delta_{ij}$. We also consider a $\mathbb{Q}$-vector space U_2 with basis denoted by $\{f_1, \ldots, f_n, f_n^*, \ldots, f_1^*\}$, with an analogous symmetric bilinear form b_2.

Let $(U, b) = (U_1, b_1) \oplus (U_2, b_2)$, let $H' = \mathrm{SO}(U, b)$ and define H to be the quotient of H' by its centre of order 2. Under the basis

$$\{e_1, \ldots, e_n, e_n^*, \ldots, e_1^*, f_1, \ldots, f_n, f_n^*, \ldots, f_1^*\},$$

the stable involution is given by conjugation with $\mathrm{diag}(1,\dots,1,-1,\dots,-1)$. Similarly to the D_{2n+1} case, we have

$$V = \left\{ \begin{pmatrix} 0 & A \\ -A^* & 0 \end{pmatrix} \,\middle|\, A \in \mathrm{Mat}_{2n\times 2n} \right\},$$

where A^* denotes reflection by the antidiagonal. In this case, the group $G = (H^\theta)^\circ$ is isomorphic to $\mathrm{SO}(U_1)\times\mathrm{SO}(U_2)/\Delta(\mu_2)$, where $\Delta(\mu_2)$ denotes the diagonal inclusion of μ_2 into the centre $\mu_2\times\mu_2$ of $\mathrm{SO}(U_1)\times\mathrm{SO}(U_2)$. As before, we can identify V with the space of $2n\times 2n$ matrices using the map

$$\begin{pmatrix} 0 & A \\ -A^* & 0 \end{pmatrix} \longmapsto A,$$

where $(g,h)\in G$ acts by $(g,h)\cdot A = gAh^{-1}$.

We consider the maximal torus T of H given by

$$\mathrm{diag}(t_1,\dots,t_n,t_n^{-1},\dots,t_1^{-1},s_1,\dots,s_n,s_n^{-1},\dots,s_1^{-1}).$$

A basis of simple roots for H and G are given by

$$\begin{aligned} S_H &= \{t_1-s_1, s_1-t_2,\dots,s_{n-1}-t_n, t_n-s_n, s_n+t_n\},\\ S_G &= \{t_1-t_2,\dots,t_{n-1}-t_n,t_{n-1}+t_n\}\\ &\quad\cup\{s_1-s_2,\dots,s_{n-1}-s_n,s_{n-1}+s_n\}. \end{aligned}$$

Let $\alpha_i = t_i/t_{i+1}$ and $\gamma_i = s_i/s_{i+1}$ for $i=1,\dots,n$, and let $\alpha_n = t_{n-1}t_n$ and $\gamma_n = s_{n-1}s_n$. Under this change of variables, the volume of $W_\flat$ is:

$$X^{2n^2}\left(\prod_{i=1}^{n-2}\alpha_i^{-2in+i^2-i}\alpha_{n-1}^{\frac{-n^2-n+4}{2}}\alpha_n^{\frac{-n^2-n}{2}}\prod_{i=1}^{n-2}\gamma_i^{-2in+i^2+i}(\gamma_{n-1}\gamma_n)^{\frac{-n^2+n}{2}}\right).$$

The modular function is

$$\delta^{-1}(t) = \prod_{i=1}^{n-2}(\alpha_i\gamma_i)^{i^2-2in+i}(t)(\alpha_{n-1}\gamma_{n-1}\alpha_n\gamma_n)^{-(n-1)n/2}(t).$$

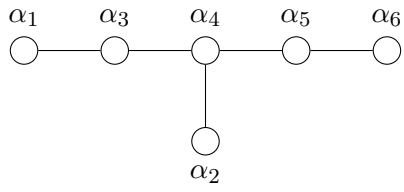
As before, we can compute:

$$\prod_{v\in W_\flat} Xw(v)\delta^{-1}(t) = X^{4n^2}\frac{Z(\beta)}{Z(b)},$$

where $Z(\beta) = \prod_{i=1}^{n-1}(\beta_{2i-1}\beta_{2i})^{2i}\cdot(\beta_{2n-1}\beta_{2n})^n$.

5.6.3. *E_6.* For the E_6 case, we use the conventions and computations in [38, Sections 2.3 and 5].

Let $S_H = \{\alpha_1, \ldots, \alpha_6\}$, where the Dynkin diagram of H is:



The pinned automorphism ϑ consists of a reflection around the vertical axis. We can define a root basis $S_G = \{\gamma_1, \gamma_2, \gamma_3, \gamma_4\}$ of G as $\gamma_1 = \alpha_3 + \alpha_4$, $\gamma_2 = \alpha_1$, $\gamma_3 = \alpha_3$ and $\gamma_4 = \alpha_2 + \alpha_4$. Under this basis, we have

$$\prod_{v \in W_\flat} X\omega(v) = X^{22}(\gamma_1^{-12}\gamma_2^{-18}\gamma_3^{-22}\gamma_4^{-12})$$

The modular function is

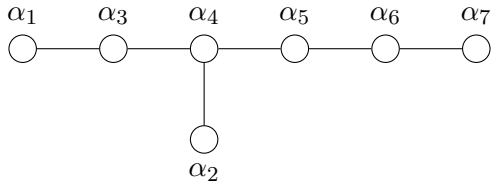
$$\delta^{-1}(t) = \left(\gamma_1^8\gamma_2^{14}\gamma_3^{18}\gamma_4^{10}\right)(t).$$

The weights of the height-one coefficients are $\{\gamma_2, -\gamma_1+\gamma_3+\gamma_4, \gamma_3, \gamma_1-\gamma_3\}$. In light of this, we obtain

$$\prod_{v \in W_\flat} X\omega(v)\delta^{-1}(t) = X^{42}\frac{Z(\beta)}{Z(b)}.$$

where $Z(\beta) = \beta_1^4\beta_2^2\beta_3^8\beta_4^6$.

5.6.4. ***E_7.*** For the E_7 and E_8 cases, we follow the conventions in [28]. Let $S_H = \{\alpha_1, \ldots, \alpha_7\}$, where the Dynkin diagram of H is:



The root basis $S_G = \{\gamma_1, \ldots, \gamma_7\}$ can be described as

$$\begin{aligned}
\gamma_1 &= \alpha_3 + \alpha_4 \\
\gamma_2 &= \alpha_5 + \alpha_6 \\
\gamma_3 &= \alpha_2 + \alpha_4 \\
\gamma_4 &= \alpha_1 + \alpha_3 \\
\gamma_5 &= \alpha_4 + \alpha_5 \\
\gamma_6 &= \alpha_6 + \alpha_7 \\
\gamma_7 &= \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5
\end{aligned}$$

The volume of $W_\flat$ can be computed to be

$$\prod_{v\in W_\flat} Xw(v) = X^{35}(\gamma_1^{-15/2}\gamma_2^{-13}\gamma_3^{-33/2}\gamma_4^{-18}\gamma_5^{-35/2}\gamma_6^{-15}\gamma_7^{-21/2}).$$

The modular function for G can be computed to be

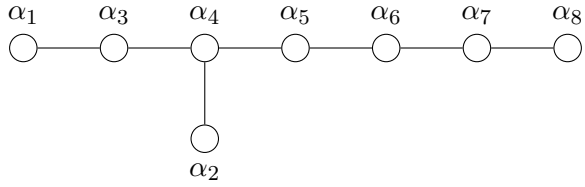
$$\delta^{-1}(t) = (\gamma_1^7\gamma_2^{12}\gamma_3^{15}\gamma_4^{16}\gamma_5^{15}\gamma_6^{12}\gamma_7^7)(t).$$

We can compute the weights β_i corresponding to the height-one coefficients, with the end result being

$$\prod_{v\in W_\flat} Xw(v)\delta^{-1}(t) = X^{70}\frac{Z(\beta)}{Z(b)},$$

for $Z(\beta) = \beta_1^2\beta_2^5\beta_3^6\beta_4^8\beta_5^7\beta_6^4\beta_7^3$.

5.6.5. ***E*$_8$.** Let $S_H = \{\alpha_1, \ldots, \alpha_8\}$, where the Dynkin diagram of H is:



The root basis $S_G = \{\gamma_1, \ldots, \gamma_8\}$ can be described as

$$\begin{aligned}
\gamma_1 &= \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5\\
\gamma_2 &= \alpha_6 + \alpha_7\\
\gamma_3 &= \alpha_4 + \alpha_5\\
\gamma_4 &= \alpha_1 + \alpha_3\\
\gamma_5 &= \alpha_2 + \alpha_4\\
\gamma_6 &= \alpha_5 + \alpha_6\\
\gamma_7 &= \alpha_7 + \alpha_8\\
\gamma_8 &= \alpha_3 + \alpha_4
\end{aligned}$$

The volume of $W_\flat$ can be computed to be

$$\prod_{\omega\in W_\flat} X\omega(t) = X^{64}(\gamma_1^{-18}\gamma_2^{-30}\gamma_3^{-40}\gamma_4^{-48}\gamma_5^{-54}\gamma_6^{-58}\gamma_7^{-30}\gamma_8^{-30}).$$

The modular function for G can be computed to be

$$\delta^{-1}(t) = (\gamma_1^{14}\gamma_2^{26}\gamma_3^{36}\gamma_4^{44}\gamma_5^{50}\gamma_6^{54}\gamma_7^{28}\gamma_8^{28})(t).$$

We get

$$\prod_{v\in W_\flat} Xw(v)\delta^{-1}(t) = X^{128}\frac{Z(\beta)}{Z(b)},$$

with $Z(\beta) = \beta_1^4\beta_2^8\beta_3^{10}\beta_4^{14}\beta_5^{12}\beta_6^8\beta_7^6\beta_8^2$.

6. Proof of the main results

We are finally in a position to prove Theorems 1.1 and 1.2. Before that, we present an auxiliary result bounding the elements in $V(\mathbb{Z})$ with big stabiliser.

6.1. Congruence conditions. We want to bound the number of elements in $V(\mathbb{Z})$ having a big stabiliser in the cusp. To do that, we will apply the Selberg sieve, which in turn requires a power saving estimate in the count of the elements in the cusp when applying finitely many congruence conditions.

Let $S \subset V(\mathbb{Z})$ be a subset which is not necessarily $G(\mathbb{Z})$-invariant. Analogously to Section 5.1, we define

$$N^{cusp}(S,X) = \sum_{i=1}^{k} \frac{1}{r_i \operatorname{vol}(G_0)} \int_{g\in\mathcal{F}} \#\{v \in S \cap W_0 \cap (gG_0\kappa(\Lambda L_i))_{<X}\}\mathrm{d}g.$$

This is the analogue of the definition of $N(S,X)$ but substituting $V(\mathbb{Z})^{red}$ for the cusp $W_0(\mathbb{Z})$. In the proof of Theorem 5.1 we saw that

$$N^{cusp}(W_0(\mathbb{Z}),X) = CX^{\dim V} + O(X^{\dim V-1})$$

for some constant C. The main theorem of this section is the following:

Theorem 6.1. *Let S be a translate of $mV(\mathbb{Z})$, for some integer $m \geq 1$. Then, for a fixed m, we have that*

$$N^{cusp}(S,X) = Cm^{-\dim V}X^{\dim V} + O(m^{1-\dim V}X^{\dim V-1}),$$

where the implied constant is independent of m and the choice of translate S, as long as $m = O(X)$.

Proof. The computation is almost exactly the same as in Section 5.1, with the only major difference being in the application of Davenport's lemma. In our situation, given a bounded region $\mathcal{R}$ as in the statement of Proposition 5.2, we have that

$$\#(\mathcal{R} \cap \mathbb{Z}^n) = \operatorname{vol}(\mathcal{R}) + O(\max\{\operatorname{vol}(\overline{\mathcal{R}}), 1\}).$$

If we now replace $\mathbb{Z}^n$ by a translate L of $m\mathbb{Z}^n$, we can translate and shrink the region $\mathcal{R}$ appropriately so that L gets identified with $\mathbb{Z}^n$, so that Davenport's lemma yields

$$\#(\mathcal{R} \cap L) = m^{-n}\operatorname{vol}(\mathcal{R}) + O(\max\{\operatorname{vol}(\overline{m^{-1}\cdot\mathcal{R}}), 1\}).$$

In our situation, what we get now instead of (5.5) is

$$\#(S \cap (nt\mathcal{B}_X)|_b) = m^{-\dim V}\operatorname{vol}((nt\mathcal{B}_X)|_b) + O(m^{1-\dim V}X^{\dim V-1}),$$

where the implied constant does not change with respect to m or S. The hypothesis that $m = O(X)$ guarantees that none of the lower-dimensional

terms dominate. Now, the rest of the argument of Section 5.4 goes through in an analogous way to obtain the desired result. □

Remark 6.2. The added hypothesis of $m = O(X)$ is added here for convenience, and does not affect the use of the Selberg sieve. In the notation of Section 5.2, we are really only adding the error terms for $d < D_0$, where D_0 is later chosen to be a suitable power of X. We can always impose the additional restriction that $D_0 = O(X)$, and the argument would go through as usual just with a possibly worse error term. We do not do the explicit computations of D_0 in this paper, but they always turn out to be $O(X)$.

6.2. Elements with big stabiliser. Let $V^{bs}(\mathbb{Z})$ be the set of elements $v \in V(\mathbb{Z})$ with $\#\operatorname{Stab}_{G(\mathbb{Q})}(v) > 1$. Then, we are in a position to prove the following:

Proposition 6.3. *There exists a constant $\delta_{bs} > 0$ such that*

$$N^{cusp}(V^{bs}(\mathbb{Z}), X) = O(X^{\dim V - \delta_{bs}}).$$

Proof. By [21, Proof of Lemma 8.22], the density of elements in $V(\mathbb{F}_p)$ having big stabiliser converges to a constant $c \in (0,1)$ as $p \to \infty$. The proof can be easily modified to show this is also true when substituting $V(\mathbb{F}_p)$ by $W_0(\mathbb{F}_p)$. Then, we can apply the Selberg sieve as explained Section 5.2, combined with Theorem 6.1. □

Remark 6.4. We remark that this result depends on Theorem 6.1. Namely, to apply the Selberg sieve in that way we need a power saving estimate on the count of reducible $G(\mathbb{Z})$-orbits in $B(\mathbb{Z})$, so we could not have proven Proposition 6.3 at the same time as Proposition 5.3.

6.3. Elements with large Q-invariant. In this section, we conclude the proof of Theorem 1.2 about bounding elements with discriminant divisible by the square of a large squarefree number. For $\mathcal{W}_m^{(1)}$, the strongly divisible case, it suffices to use the Ekedahl sieve as in [2, Theorem 3.5, Lemma 3.6], knowing that the discriminant polynomial is irreducible by [21, Lemma 4.2]. Thus, to conclude the proof of Theorem 1.2, it suffices to consider the weakly divisible case.

Recall that by the results in Section 3, to prove Theorem 1.2 it is enough to bound the number of elements in

$$W_M = \left\{ v \in \frac{1}{N} V(\mathbb{Z}) \;\middle|\; \begin{array}{l} v = g\kappa_b \text{ for a squarefree } m > M,\ (m,N) = 1, \\ g \in G(\mathbb{Z}[1/m]) \setminus G(\mathbb{Z}),\ b \in B(\mathbb{Z}), \Delta(b) \neq 0 \end{array} \right\}.$$

It suffices to prove that:

Theorem 6.5. *There exists a constant $\delta > 0$ such that*

$$N(W_M, X) = O\left(\frac{X^{\dim V}}{M}\right) + O(X^{\dim V - \delta}).$$

Proof. We can apply the same averaging argument as in Section 5.1, where we get that $N(W_{M,i}, X)$ is equal, up to an error of the order of $O(X^{\dim V-\delta})$, to

$$\frac{1}{r_i \operatorname{vol}(G_0)} \int_{n\in\omega} \int_{t\in T_c} \#\{v \in W_M \cap W_0(\mathbb{Z}) \cap (ntG_0\kappa(\Lambda L_i))\}\delta^{-1}(t)\mathrm{d}n\mathrm{d}^\times t.$$

Here, $W_{M,i} := W_M \cap G(\mathbb{R})\kappa(\Lambda L_i)$. In light of Proposition 6.3, it suffices to count elements in $W_{M,i} \cap W_0(\mathbb{Z})$ with trivial stabiliser. For each of these elements v, Proposition 3.1 guarantees that $Q(v) > M$, and in particular that $|Z(v)| > M^2$. Then, by following the same proof as in Theorem 5.1, there is some constant C_i such that

$$N(W_{M,i}, X) = C_i X^{\dim V} \sum_{\substack{b\in\mathbb{Z}^r \\ |Z(b)|>M^2}} \frac{1}{|Z(b)|} + O(X^{\dim V-\delta}).$$

But the written sum is $O(\frac{1}{M})$, so that concludes the proof. □

Therefore, we have proven Theorem 1.2. With the same proof as [7, Theorem 4.4], combining the estimates for the strongly divisible primes and the weakly divisible primes, we get:

Theorem 6.6. *For a squarefree integer m, let $\mathcal{W}_m$ denote the elements of $B(\mathbb{Z})$ with discriminant divisible by m^2. There is a constant $\delta > 0$ such that*

$$\sum_{\substack{m>M \\ m \text{ squarefree} \\ (m,N)=1}} \#\{b \in \mathcal{W}_m \mid \mathrm{ht}(b) < X\} = O_\varepsilon\left(\frac{X^{\dim V+\varepsilon}}{\sqrt{M}}\right) + O(X^{\dim V-\delta}).$$

6.4. A squarefree sieve. Theorem 1.1 follows from the previous tail estimates by performing a squarefree sieve, following the methods in [7, Section 4]. In fact, we will prove a slightly more general result about counting elements in $B(\mathbb{Z})$ imposing infinitely many congruence conditions.

Let κ be a positive integer. We say a subset $\mathcal{S} \subset B(\mathbb{Z})$ is *κ-acceptable* if $\mathcal{S} = B(\mathbb{Z}) \cap \bigcap_p \mathcal{S}_p$, where $\mathcal{S}_p \subset B(\mathbb{Z}_p)$ satisfy the following:

(1) $\mathcal{S}_p$ is defined by congruence conditions modulo p^κ.
(2) For all sufficiently large primes p, the set $\mathcal{S}_p$ contains all $b \in B(\mathbb{Z}_p)$ such that $p^2 \nmid \Delta(b)$.

For any subset $A \subset B(\mathbb{Z})$, denote by $N(A, X)$ the number of elements of A having height less than X. For any prime p and any subset $A_p \subset B(\mathbb{Z}_p)$, we denote by $\rho(A_p)$ the density of elements of A_p inside $B(\mathbb{Z}_p)$.

Theorem 6.7. *Let κ be a positive integer, and let $\mathcal{S} \subset B(\mathbb{Z})$ be a κ-acceptable subset. Then, there exists a constant $\delta > 0$ such that*

$$N(\mathcal{S}, X) = \left(\prod_p \rho(\mathcal{S}_p)\right) N(B(\mathbb{Z}), X) + O(X^{\dim V-\delta}).$$

Proof. Recall that $B = \operatorname{Spec} \mathbb{Z}[p_{d_1}, \ldots, p_{d_k}]$. For an element $b \in B(\mathbb{Z})$ of height at most X, it holds that $|p_{d_i}(b)| < X^{d_i}$, where by Table 2.2 we see that $d_i \geq 2$ for all i. For a positive squarefree integer m coprime to N, denote by $\mathcal{S}'_m$ the big family defined for each prime p as:

- If $p \mid N$, we set $\mathcal{S}'_p = \mathcal{S}_p$.
- If $p \mid m$, we set $\mathcal{S}'_p = B(\mathbb{Z}_p) \setminus \mathcal{S}_p$.
- Otherwise, we set $\mathcal{S}'_p = B(\mathbb{Z}_p)$.

By the inclusion-exclusion principle, we get that

$$N(\mathcal{S}, X) = \sum_{\substack{m \geq 1 \\ (m,N)=1}} \mu(m) N(\mathcal{S}'_m, X),$$

where $\mu(m)$ is the Möbius function. We can estimate $N(\mathcal{S}_m, X)$ as follows: in $B(\mathbb{Z})$, the set $\mathcal{S}'_m$ is the union of T_m translates of a congruence class modulo $(mN)^\kappa$, and we have that $T_m = \prod_{p|m}(1 - \rho(\mathcal{S}_p)) \prod_{p|N} \rho(\mathcal{S}_p) \cdot (mN)^{k\kappa}$. Each of these congruence classes contributes $\prod_{i=1}^k \left(\frac{2X^{d_i}}{(mN)^\kappa} + O(1) \right)$ to the sum $N(\mathcal{S}_m, X)$. In summary, we get

$$N(\mathcal{S}_m, X) = \prod_{p|m}(1 - \rho(\mathcal{S}_p)) \prod_{p|N} \rho(\mathcal{S}_p) N(B(\mathbb{Z}), X) + O(m^\kappa X^{\dim V - 2}).$$

By Theorem 1.2, we also have that for large enough M:

$$\sum_{m \geq M} \mu(m) N(\mathcal{S}'_m, X) = O_\varepsilon \left(\frac{X^{\dim V + \varepsilon}}{\sqrt{M}} \right) + O(X^{\dim V - \delta})$$

Combining the previous identities, we get

$$\begin{aligned} N(\mathcal{S}, X) &= \prod_{p|N} \rho(\mathcal{S}_p) \sum_{m=1}^{M} \mu(m) \prod_{p|m} (1 - \rho(\mathcal{S}_p)) N(B(\mathbb{Z}), X) \\ &\quad + O_\varepsilon \left(M^{\kappa+1} X^{\dim V - 2} + \frac{X^{\dim V + \varepsilon}}{\sqrt{M}} + X^{\dim V - \delta} \right) \\ &= \prod_{p} \rho(\mathcal{S}_p) N(B(\mathbb{Z}), X) \\ &\quad + O_\varepsilon \left(\frac{X^{\dim V}}{M} + M^{\kappa+1} X^{\dim V - 2} + \frac{X^{\dim V + \varepsilon}}{\sqrt{M}} + X^{\dim V - \delta} \right), \end{aligned}$$

where the last estimate follows from the observation that $\rho(\mathcal{S}_p) \gg 1 - \frac{1}{p^2}$ by [26, Proof of Theorem 3.2]. Now, optimising we choose $M = X^{4/(2\kappa+3)}$, which is enough for the result. □

References

[1] J. D. Adler, J. Fintzen & S. Varma, "On Kostant sections and topological nilpotence", *J. Lond. Math. Soc. (2)* **97** (2018), no. 2, p. 325-351.

[2] M. Bhargava, "The geometric sieve and the density of squarefree values of invariant polynomials", 2014, https://arxiv.org/abs/1402.0031.

[3] ———, "Rational points on elliptic and hyperelliptic curves", in *Proceedings of the International Congress of Mathematicians – Seoul 2014. Vol. I: Plenary lectures and ceremonies*, KM Kyung Moon Sa, 2014, p. 657-684.

[4] M. Bhargava & 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*, American Mathematical Society, 2013, p. 23-91.

[5] M. Bhargava & W. Ho, "On average sizes of Selmer groups and ranks in families of elliptic curves having marked points", 2022, https://arxiv.org/abs/2207.03309.

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

[7] M. Bhargava, A. Shankar & X. Wang, "Squarefree values of polynomial discriminants I", *Invent. Math.* **228** (2022), no. 3, p. 1037-1073.

[8] ———, "Squarefree values of polynomial discriminants II", 2022, https://arxiv.org/abs/2207.05592.

[9] A. Borel, "Density and maximality of arithmetic subgroups.", *J. Reine Angew. Math.* **1966** (1966), no. 224, p. 78-89.

[10] A. Borel & J.-P. Serre, "Corners and arithmetic groups", *Comment. Math. Helv.* **48** (1973), p. 436-491.

[11] A. J. Coleman, "The Betti Numbers of the Simple Lie Groups", *Can. J. Math.* **10** (1958), p. 349-356.

[12] B. Conrad, "Reductive group schemes", in *Autour des schémas en groupes. Vol. I*, Panoramas et Synthèses, vol. 42/43, Société Mathématique de France, 2014, p. 93-444.

[13] H. Davenport, "On a principle of Lipschitz", *J. Lond. Math. Soc.* **26** (1951), p. 179-183.

[14] H. Esnault, "Sur l'identification de singularités apparaissant dans des groupes algébriques complexes", in *Seminar on Singularities, Paris, 1976 / 1977*, Publications Mathématiques de l'Université Paris VII, vol. 7, 1980, p. 31-59.

[15] A. Grothendieck, "Éléments de géométrie algébrique : IV. Étude locale des schémas et des morphismes de schémas, Quatrième partie", *Publ. Math., Inst. Hautes Étud. Sci.* **32** (1967), p. 5-361.

[16] W. Ho, "How many rational points does a random curve have?", *Bull. Am. Math. Soc.* **51** (2013), no. 1, p. 27-52.

[17] H. Iwaniec & E. Kowalski, *Analytic number theory*, Colloquium Publications, vol. 53, American Mathematical Society, 2004, xii+615 pages.

[18] J. Jantzen, *Representations of algebraic groups*, Mathematical Surveys and Monographs, American Mathematical Society, 2007.

[19] A. W. Knapp, *Lie groups beyond an introduction*, Progress in Mathematics, vol. 140, Birkhäuser, 1996, xvi+604 pages.

[20] J. Laga, "The average size of the 2-Selmer group of a family of non-hyperelliptic curves of genus 3", *Algebra Number Theory* **16** (2022), no. 5, p. 1161-1212.

[21] ———, "Graded Lie Algebras, Compactified Jacobians and Arithmetic Statistics", *J. Eur. Math. Soc.* (2024), Published online first.

[22] S. Lang, $SL_2(\mathbf{R})$, Addison-Wesley Publishing Group, 1975.

[23] R. P. Langlands, "The volume of the fundamental domain for some arithmetical subgroups of Chevalley groups", in *Algebraic Groups and Discontinuous Subgroups (Proc. Sympos. Pure Math., Boulder, Colo., 1965)*, American Mathematical Society, 1966, p. 143-148.

[24] D. Panyushev, "On invariant theory of θ-groups", *J. Algebra* **283** (2005), p. 655-670.

[25] V. Platonov & A. Rapinchuk, *Algebraic Groups and Number Theory*, Pure and Applied Mathematics, vol. 139, Academic Press Inc., 1994, Translated from the 1991 Russian original by Rachel Rowen.
[26] B. Poonen, "Squarefree values of multivariable polynomials", *Duke Math. J.* **118** (2003), no. 2, p. 353-373.
[27] R. W. Richardson, "Orbits, invariants, and representations associated to involutions of reductive groups", *Invent. Math.* **66** (1982), no. 2, p. 287-312.
[28] B. Romano & J. Thorne, "On the arithmetic of simple singularities of type E", *Res. Number Theory* **4** (2018), no. 2, article no. 21 (34 pages).
[29] ———, "E_8 and the average size of the 3-Selmer group of the Jacobian of a pointed genus-2 curve", *Proc. Lond. Math. Soc. (3)* **122** (2021), no. 5, p. 678-723.
[30] A. N. Shankar, "2-Selmer groups of hyperelliptic curves with marked points", *Trans. Am. Math. Soc.* **372** (2018), no. 1, p. 267-304.
[31] A. Shankar, A. Siad, A. Swaminathan & I. Varma, "Geometry-of-numbers methods in the cusp", 2022, https://arxiv.org/abs/2110.09466.
[32] A. Shankar & J. Tsimerman, "Counting S_5 fields with a power saving error term", *Forum Math. Sigma* **2** (2014), article no. e13 (8 pages).
[33] A. Shankar & X. Wang, "Rational points on hyperelliptic curves having a marked non-Weierstrass point", *Compos. Math.* **154** (2018), no. 1, p. 188-222.
[34] P. Slodowy, *Simple singularities and simple algebraic groups*, Lecture Notes in Mathematics, Springer, 1980, x+175 pages.
[35] R. Steinberg, "Torsion in reductive groups", *Adv. Math.* **15** (1975), no. 1, p. 63-92.
[36] A. V. Sutherland, "A database of nonhyperelliptic genus-3 curves over $\mathbb{Q}$", in *ANTS XIII. Proceedings of the thirteenth algorithmic number theory symposium, University of Wisconsin-Madison, WI, USA, July 16–20, 2018*, The Open Book Series, vol. 2, Mathematical Sciences Publishers, 2019, p. 443-459.
[37] J. Thorne, "Vinberg's representations and arithmetic invariant theory", *Algebra Number Theory* **7** (2013), no. 9, p. 2331-2368.
[38] ———, "E_6 and the arithmetic of a family of non-hyperelliptic curves of genus 3", *Forum Math. Pi* **3** (2015), article no. e1 (41 pages).
[39] ———, "Arithmetic invariant theory and 2-descent for plane quartic curves", *Algebra Number Theory* **10** (2016), no. 7, p. 1373-1413.

Martí Oller
University of Cambridge, Department of Pure Mathematics and Mathematical Statistics, Centre for Mathematical Sciences, Wilberforce Road,Cambridge CB3 0WB, United Kingdom
E-mail: mo512@cam.ac.uk
URL: https://www.dpmms.cam.ac.uk/~mo512/