Journal de l'École polytechnique
Mathématiques

Christopher Daw & Martin Orr

Lattices with skew-Hermitian forms over division algebras and unlikely intersections

Tome 10 (2023), p. 1097-1156.

https://doi.org/10.5802/jep.240

Publié avec le soutien
du Centre National de la Recherche Scientifique

Publication membre du
Centre Mersenne pour l'édition scientifique ouverte
www.centre-mersenne.org
e-ISSN : 2270-518X

Journal de l'École polytechnique
Mathématiques

Tome 10, 2023, p. 1097–1156 DOI: 10.5802/jep.240

LATTICES WITH SKEW-HERMITIAN FORMS OVER DIVISION ALGEBRAS AND UNLIKELY INTERSECTIONS

by Christopher Daw & Martin Orr

Abstract. — This paper has two objectives. First, we study lattices with skew-Hermitian forms over division algebras with positive involutions. For division algebras of Albert types I and II, we show that such a lattice contains an "orthogonal" basis for a sublattice of effectively bounded index. Second, we apply this result to obtain new results in the field of unlikely intersections. More specifically, we prove the Zilber–Pink conjecture for the intersection of curves with special subvarieties of simple PEL type I and II under a large Galois orbits conjecture. We also prove this Galois orbits conjecture for certain cases of type II.

Résumé (Réseaux munis de formes anti-hermitiennes sur des algèbres à division et intersections atypiques)

Cet article a deux objectifs. Nous étudions d'abord les réseaux munis de formes anti-hermitiennes sur des algèbres à division avec involutions positives. Pour les algèbres à division de type I et II dans la classification d'Albert, nous montrons qu'un tel réseau contient une base « orthogonale » pour un sous-réseau d'indice borné de manière effective. Ensuite, nous appliquons ce résultat pour obtenir de nouveaux résultats dans la théorie des intersections atypiques. En particulier, nous prouvons la conjecture de Zilber–Pink pour l'intersection de courbes avec les sous-variétés spéciales de type PEL simple I et II en supposant la conjecture des grandes orbites galoisiennes vraie. De plus, nous prouvons cette conjecture sur les orbites galoisiennes dans certains cas de type II.

Contents

Mathematical subject classification (2020). — 11E39, 11G18.
Keywords. — Division algebras, Hermitian forms, abelian varieties, Zilber–Pink conjecture, unlikely intersections.

This work was supported by the Engineering and Physical Sciences Research Council [EP/S029613/1 to C.D., EP/T010134/1 to M.O.].

1. Introduction

In this paper we develop a quantitative result on reduction theory for lattices over division algebras equipped with skew-Hermitian forms. Our main theorem is inspired by Minkowski's theorems on lattices and Masser and Wüstholz's class index lemma [MW95], with the additional ingredient of looking for a basis which interacts nicely with a skew-Hermitian form.

Our purpose in proving this theorem is to apply it to certain cases of the Zilber–Pink conjecture in moduli spaces of abelian varieties. The theorem on lattices supplies the "parameter height bound" needed for the Pila–Zannier strategy. This generalises our earlier paper [DO22], where we proved some cases of Zilber–Pink for the moduli space of abelian *surfaces* using quantitative reduction theory.

1.A. Bases and skew-Hermitian forms over division algebras. —

A classical result in algebraic number theory, due to Minkowski, asserts that if R is the ring of integers of a number field, then every ideal $I \subset R$ contains an element x such that the index $[I : Rx]$ is bounded by an explicit multiple of $\sqrt{\operatorname{disc}(R)}$. A similar result can be proved for torsion-free modules of finite rank over the ring of integers of a number field, by combining Minkowski's theorem with the structure theory of finite-rank modules over a Dedekind domain (see [CR62, §22, Ex. 6]).

In [MW95], Masser and Wüstholz generalised this theorem to torsion-free R-modules L of finite rank over any order R in a division $\mathbb{Q}$-algebra. This generalisation shows that there is a free R-submodule of finite index in L, with index $[L : R]$ bounded polynomially in terms of $\operatorname{disc}(R)$. The statement is as follows. (See section 2.F for the definition of the discriminant of an order in a semisimple $\mathbb{Q}$-algebra.)

Theorem 1.1 ([MW95, Chap. 2, Class Index Lem.]) — *Let D be a division $\mathbb{Q}$-algebra and let R be an order in D. Let L be a torsion-free R-module of finite rank m. Then there exists a left D-basis $v_1, \dots, v_m$ for $D \otimes_R L$ such that $v_1, \dots, v_m$ are in L and $[L : Rv_1 + \cdots + Rv_m] \leqslant |\operatorname{disc}(R)|^{m/2}$.*

In another direction, if L is a $\mathbb{Z}$-module of finite rank equipped with a positive definite symmetric bilinear form $\psi\colon L \times L \to \mathbb{Z}$, then one can use the classical reduction theory of quadratic forms to find an *orthogonal* basis $v_1, \dots, v_m$ for $L \otimes_{\mathbb{Z}} \mathbb{Q}$ such that $v_1, \dots, v_n \in L$ and $[L : \mathbb{Z}v_1 + \cdots + \mathbb{Z}v_m]$ is bounded by a polynomial in $|\operatorname{disc}(L)|$. A similar result for a $\mathbb{Z}$-module of finite rank equipped with a symplectic form can be found in [Orr15] (see Lemma 4.3 therein).

In this paper, we obtain a version of Theorem 1.1 in which L is equipped with a $(D, \dagger)$-skew-Hermitian form (see section 3.A for the definition of a $(D, \dagger)$-skew-Hermitian form.) We seek a basis of $D \otimes_R L$ which is weakly symplectic or weakly unitary with respect to this form. Weakly symplectic or weakly unitary bases are the analogues for $(D, \dagger)$-skew-Hermitian forms of bases which are orthogonal but not necessarily orthonormal: we say that a D-basis $v_1, \dots, v_m$ is *weakly symplectic* if

$\psi(v_i, v_j) = 0$ for all i, j except when $\{i, j\} = \{2k-1, 2k\}$ for some $k \in \mathbb{Z}$, and that the basis is *weakly unitary* if $\psi(v_i, v_j) = 0$ for all $i, j \in \{1, \dots, m\}$ such that $i \neq j$.

Theorem 1.2. — *Let D be either a totally real number field or a totally indefinite quaternion algebra over a totally real number field. Let $\dagger$ be a positive involution of D. Let V be a left D-vector space of dimension m, equipped with a non-degenerate $(D, \dagger)$-skew-Hermitian form $\psi\colon V \times V \to D$. Let L be a $\mathbb{Z}$-lattice of full rank in V such that $\mathrm{Trd}_{D/\mathbb{Q}}\, \psi(L \times L) \subset \mathbb{Z}$. Let $R = \mathrm{Stab}_D(L)$ denote the stabiliser of L in D.*

Then there exists a D-basis $v_1, \dots, v_m$ for V such that:

(i) *$v_1, \dots, v_m \in L$;*

(ii) *the basis is weakly symplectic (when D is a field) or weakly unitary (when D is a quaternion algebra) with respect to ψ;*

(iii) *$[L : Rv_1 + \cdots + Rv_m] \leqslant C_1 |\mathrm{disc}(R)|^{C_2} |\mathrm{disc}(L)|^{C_3}$;*

(iv) *$|\psi(v_i, v_j)|_D \leqslant C_4 |\mathrm{disc}(R)|^{C_5} |\mathrm{disc}(L)|^{C_6}$ for $1 \leqslant i, j \leqslant m$.*

The constants $C_1, \dots, C_6$ depend only on m and $\dim_{\mathbb{Q}}(D)$.

Explicit, but not optimal, values for the constants are given in Proposition 4.5. One could also prove a version of this theorem that bounds the lengths of the vectors v_i, in the style of Minkowski's second theorem, but this is stronger than needed for our application, and according to the proof that we know, the constants are exponential instead of polynomial in m.

Division $\mathbb{Q}$-algebras with positive involution were classified by Albert into four types (see section 2.B). The division algebras treated in Theorem 1.2 are those of types I and II in Albert's classification. It is likely that this paper's strategy could be adapted to prove Theorem 1.2 for division $\mathbb{Q}$-algebras with positive involution of types III and IV, as well as a version for Hermitian forms instead of skew-Hermitian forms, although various steps in the argument would require modification.

1.B. Applications to the Zilber–Pink conjecture. —

We apply Theorem 1.2 to prove certain cases of the Zilber–Pink conjecture on unlikely intersections in the moduli space $\mathcal{A}_g$ of principally polarised abelian varieties of dimension g (which is an example of a Shimura variety), as follows.

Theorem 1.3. — *Let $g \geqslant 3$. Let Σ denote the set of points $s \in \mathcal{A}_g(\mathbb{C})$ for which the endomorphism algebra of the associated abelian variety A_s is either a totally real field, other than $\mathbb{Q}$, or a non-split totally indefinite quaternion algebra over a totally real field. Let C be an irreducible Hodge generic algebraic curve in $\mathcal{A}_g$.*

If C satisfies Conjecture 1.4*, then $C \cap \Sigma$ is finite.*

Throughout this paper, whenever we refer to endomorphisms of an abelian variety, we refer to its endomorphisms over an algebraically closed field.

The analogous statement to Theorem 1.3 for $g = 2$ was proved in our earlier work [DO22]. In that paper, [DO22, Lem. 5.7] played the role which is now played by Theorem 1.2. Indeed this paper represents the next stage of our programme on the

Zilber–Pink conjecture for Shimura varieties, following on from [DO21] and [DO22], which were inspired by the earlier papers [HP12], [HP16], [DR18], [Orr21].

Conjecture 1.4, referred to in Theorem 1.3, is a large Galois orbits conjecture, of the sort appearing in many works on unlikely intersections (for example, [Ull14, Conj. 2.7], [HP16, Conj. 8.2], [DR18, Conj. 11.1], [DO22, Conj. 6.2]).

Conjecture 1.4. — *Define $\Sigma \subset \mathcal{A}_g$ as in Theorem* 1.3 *and let $C \subset \mathcal{A}_g$ denote an irreducible Hodge generic algebraic curve defined over a finitely generated field $L \subset \mathbb{C}$. Then there exist positive constants C_7 and C_8, depending only on g, L and C, such that, for any point $s \in C \cap \Sigma$,*

$$\#\operatorname{Aut}(\mathbb{C}/L) \cdot s \geqslant C_7 |\operatorname{disc}(\operatorname{End}(A_s))|^{C_8}.$$

The most general conjecture of this type in the context of Shimura varieties which has been written down is [DR18, Conj. 11.1]. It is not clear whether [DR18, Conj. 11.1] implies Conjecture 1.4, because it is not clear how $|\operatorname{disc}(\operatorname{End}(A_s))|$ is related to the complexity $\Delta(\langle s \rangle)$ defined in [DR18]. For example, in [DR18, Conj. 11.1], $\Delta(\langle s \rangle)$ is the complexity of the smallest special subvariety of $\mathcal{A}_g$ containing s. In Conjecture 1.4, $|\operatorname{disc}(\operatorname{End}(A_s))|$ is a measure of the complexity of the smallest special subvariety *of PEL type* containing s, which might not be the same as the smallest special subvariety containing s. However, for the purpose of proving cases of the Zilber–Pink conjecture, the precise definition of complexity is not important: we only need a parameter height bound and a Galois orbits bound which involve the same notion of complexity. Since we are focusing on special subvarieties of PEL type, the discriminant of the endomorphism ring is a natural measure of complexity.

Using André's G-functions method [And89], in the form of [DO21, Th. 8.2], we prove Conjecture 1.4 in certain cases and thus establish Theorem 1.3 unconditionally in those settings. The proof of large Galois orbits in Theorem 1.5 does not involve new ideas beyond those in [DO21], [DO22]: the new contribution of this paper is in the parameter height bound.

Theorem 1.5. — *Let g be an even positive integer. Let Σ^* denote the set of points $s \in \mathcal{A}_g$ for which $\operatorname{End}(A_s) \otimes_{\mathbb{Z}} \mathbb{Q}$ is a non-split totally indefinite quaternion algebra whose centre is a totally real field of degree e such that $4e$ does not divide g.*

Let $C \subset \mathcal{A}_g$ denote an irreducible Hodge generic algebraic curve defined over a number field. Suppose that the Zariski closure of C in the Baily–Borel compactification of $\mathcal{A}_g$ intersects the 0-dimensional stratum.

Then C satisfies Conjecture 1.4 *for Σ^* (in the place of Σ). Hence, $C \cap \Sigma^*$ is finite.*

Compared with Conjecture 1.4, Theorem 1.5 adds two restrictions: Σ^* is defined by a smaller class of endomorphism algebras than Σ, and there is a condition on the intersection of the Zariski closure of C with the boundary of the Baily–Borel compactification. We recall that the Baily–Borel compactification of the moduli space $\mathcal{A}_g$ is naturally stratified as a disjoint union

$$\mathcal{A}_g \sqcup \mathcal{A}_{g-1} \sqcup \cdots \sqcup \mathcal{A}_1 \sqcup \mathcal{A}_0$$

of locally closed subvarieties. The zero-dimensional stratum is $\mathcal{A}_0$, which is a point. The condition that C intersects the zero-dimensional stratum is equivalent to saying that the associated family of principally polarised abelian varieties degenerates to a torus (this informal statement can be made precise as in [DO21, Th. 1.4]).

1.C. The Zilber–Pink conjecture and special subvarieties of PEL type. — Let us recall a general statement of the Zilber–Pink conjecture for Shimura varieties. A *special subvariety* of a Shimura variety S means an irreducible component of a Shimura subvariety of S. An irreducible subvariety of S is *Hodge generic* if it is not contained in any special subvariety other than a component of S itself.

Conjecture 1.6 ([Pin05, Conj. 1.3]). — *Let S be a Shimura variety and let V be an irreducible Hodge generic subvariety of S. Then the intersection of V with the special subvarieties of S having codimension greater than* $\dim V$ *is not Zariski dense in V.*

In order to relate this to Theorem 1.3, we introduce a class of special subvarieties of $\mathcal{A}_g$ which come from endomorphisms of abelian varieties. We recall that $\mathcal{A}_g$ is an irreducible algebraic variety over $\mathbb{Q}$. For any algebraically closed field k containing $\mathbb{Q}$ and any point $s \in \mathcal{A}_g(k)$, we write A_s for the principally polarised abelian variety over k (defined up to isomorphism) corresponding to the point s.

For any ring R, the set

$$\mathcal{M}_R = \{s \in \mathcal{A}_g(\mathbb{C}) : \text{there exists an injective homomorphism } R \to \mathrm{End}(A_s)\}$$

is a countable union of algebraic subvarieties of $\mathcal{A}_g$. Each irreducible component of $\mathcal{M}_R$ is a special subvariety of $\mathcal{A}_g$. We call a subvariety of $\mathcal{A}_g$ a *special subvariety of PEL type* if it is an irreducible component of $\mathcal{M}_R$ for some R.

If $R \ncong \mathbb{Z}$, then $\mathcal{M}_R$ is strictly contained in $\mathcal{A}_g$. Hence the set Σ defined in Theorem 1.3 is contained in the union of the proper special subvarieties of PEL type of $\mathcal{A}_g$. Furthermore, as we prove in Proposition 5.5, for $g \geqslant 3$, all proper special subvarieties of PEL type of $\mathcal{A}_g$ have codimension at least 2. Thus, Conjecture 1.6 predicts that the intersection $C \cap \Sigma$ of Theorem 1.3 should not be Zariski dense in the curve C, that is, it should be finite.

For each special subvariety of PEL type $S \subset \mathcal{A}_g$, there is a largest ring R such that S is a component of $\mathcal{M}_R$. We call this ring R the *generic endomorphism ring* of S, and we call $R \otimes_{\mathbb{Z}} \mathbb{Q}$ the *generic endomorphism algebra* of S. We say that a point $s \in S(\mathbb{C})$ is *endomorphism generic* if the endomorphism ring of A_s is equal to R. Note that all points in the complement of countably many proper subvarieties of S are endomorphism generic.

We call $S \subset \mathcal{A}_g$ a *special subvariety of simple PEL type* if it is a special subvariety of PEL type and its generic endomorphism algebra is a division algebra. (Equivalently, A_s is a simple abelian variety for endomorphism generic points $s \in S(\mathbb{C})$.) We call S a *special subvariety of simple PEL type I or II* if it is a special subvariety of PEL type whose generic endomorphism ring is a division algebra of type I or II in the Albert classification (see section 2.B). Thus the set Σ in Theorem 1.3 is the union of

the endomorphism generic loci of all special subvarieties of simple PEL type I or II, excluding $\mathcal{A}_g$ itself.

In section 5.C we establish the following bounds on the dimensions of special subvarieties of PEL type in $\mathcal{A}_g$. These are not necessary for proving Theorem 1.3, but they are interesting for understanding the Zilber–Pink conjecture in the context of special subvarieties of PEL type. In particular, when $g \geqslant 3$, Proposition 5.5 guarantees that intersections between a Hodge generic curve and all proper special subvarieties of PEL type in $\mathcal{A}_g$ are predicted to be "unlikely" by the Zilber–Pink conjecture.

Proposition 1.7. — *Let $S \subset \mathcal{A}_g$ be a special subvariety, not equal to $\mathcal{A}_g$.*

(i) *If S is of simple PEL type, then* $\dim(S) \leqslant \dim(\mathcal{A}_g) - g^2/4$.

(ii) *If S is of PEL type, then* $\dim(S) \leqslant \dim(\mathcal{A}_g) - g + 1$.

We also prove a finiteness result for special subvarieties of simple PEL type I or II of bounded complexity (Corollary 8.4). This is the analogue of a special case of [DR18, Conj. 10.3], using our notion of complexity (cf. discussion of complexity of special subvarieties below Conjecture 1.4).

Proposition 1.8. — *Define $\Sigma \subset \mathcal{A}_g$ as in Theorem* 1.3. *For each $b \in \mathbb{R}$, the points $s \in \Sigma$ such that $|\mathrm{disc}(\mathrm{End}(A_s))| \leqslant b$ belong to only finitely many proper special subvarieties of simple PEL type I or II.*

1.D. High-level proof strategy for Theorem 1.3. — We now outline the strategy of the proof of Theorem 1.3, which is carried out in sections 5 to 8, making use of Theorem 1.2. For our notation and terminology around Shimura datum components, see [DO21, §§2.A & 2.B].

Let $\mathbf{G} = \mathbf{GSp}_{2g}$ and let $(\mathbf{G}, X^+)$ denote the Shimura datum component defined in section 5.A, which gives rise to the Shimura variety $\mathcal{A}_g$. By Lemma 5.1, the Shimura subdatum components $(\mathbf{H}, X_{\mathbf{H}}^+)$ associated with special subvarieties of simple PEL type I or II lie in only finitely many $\mathbf{G}(\mathbb{R})$-conjugacy classes. Hence it suffices to prove Theorem 1.3 "one $\mathbf{G}(\mathbb{R})$-conjugacy class at a time." Thanks to Lemma 5.1, this means that we choose positive integers d, e, m and let $\mathbf{H}_0$ be the subgroup of $\mathbf{G} = \mathbf{GSp}_{2g}$ defined in (5.2) for these d, e, m. We prove Theorem 1.3 with Σ replaced by $\Sigma_{d,e,m}$, namely, the union of the endomorphism generic loci of all proper special subvarieties of $\mathcal{A}_g$ of simple PEL type I or II whose underlying group is $\mathbf{G}(\mathbb{R})$-conjugate to $\mathbf{H}_0$.

Let π denote the standard quotient map $X^+ \to \mathcal{A}_g(\mathbb{C})$ and let $\mathcal{F}_g$ denote a Siegel fundamental set of X^+, as defined in [Orr18, §2] and [DO21, §2.G].

In order to prove Theorem 1.3 for $\Sigma_{d,e,m}$, we follow the same proof strategy as [DO22] (which proves the analogous result for $g = 2$, $d = 2$, $e = m = 1$). The idea is to apply the Habegger–Pila–Wilkie counting theorem [HP16, Cor. 7.2] to a definable set of the form

$$D = \{(y, z) \in Y \times \mathcal{C} : z \in X_y\},$$

where $Y \subset \mathbb{R}^n$ is a semi-algebraic parameter space for subsets $X_y \subset X^+$ and $\mathcal{C} = \pi^{-1}(C(\mathbb{C})) \cap \mathcal{F}_g$. The parameter space Y has the following properties:

(1) For every rational point $y \in Y \cap \mathbb{Q}^n$, X_y is a pre-special subvariety of X^+ whose underlying group is $\mathbf{G}(\mathbb{R})$-conjugate to $\mathbf{H}_0$.

(2) For every point $s \in \Sigma_{d,e,m}$, there exists $z \in \pi^{-1}(s) \cap \mathcal{F}_g$ such that z lies in X_y for some rational point $y \in Y \cap \mathbb{Q}^n$, with the height $H(y)$ polynomially bounded in terms of $\mathrm{End}(A_s)$.

Consequently, if $C \cap \Sigma_{d,e,m}$ is infinite, and if the large Galois orbits conjecture holds, then the number of points $(y, z) \in D$ with $y \in Y \cap \mathbb{Q}^n$ grows reasonably quickly with respect to $H(y)$. Then the Habegger–Pila–Wilkie theorem tells us that D contains a path whose projection to Y is semi-algebraic and whose projection to $\mathcal{C}$ is non-constant. We can conclude by a functional transcendence argument as in [DO22, §6.5].

1.E. Proof strategy: parameter space. — The strategy described in section 1.D is the same as that applied in [HP16], [DR18], [DO21], [DO22], and others. The new ingredient required to apply the strategy described in section 1.D in our case is to construct a suitable parameter space Y for special subvarieties of simple PEL type I or II and prove that it satisfies property (2) above.

To construct Y, we will choose a suitable representation $\rho\colon \mathbf{G} \to \mathbf{GL}(W)$, where W is a $\mathbb{Q}$-vector space, and a vector $w_0 \in W$ such that $\mathrm{Stab}_{\mathbf{G}}(w_0) = \mathbf{H}_0$. Then we define Y to be the "expanded ρ-orbit" of w_0 in $W_{\mathbb{R}}$:

$$Y = \mathrm{Aut}_{\rho(\mathbf{G})}(W_{\mathbb{R}})\, \rho(\mathbf{G}(\mathbb{R}))\, w_0.$$

For each $y \in Y$, we define $\mathbf{H}_y = \mathrm{Stab}_{\mathbf{G}_{\mathbb{R}}}(y)$ and

$$X_y = \{z \in X^+ : z(\mathbb{S}) \subset \mathbf{H}_y\}.$$

If $y \in Y \cap \mathbb{Q}^n$, then $\mathbf{H}_y$ is a $\mathbb{Q}$-algebraic subgroup of $\mathbf{G}$, which is $\mathbf{G}(\mathbb{R})^+$-conjugate to $\mathbf{H}_0$. By Lemma 5.1, we have $\mathbf{H}_{y,\mathbb{R}} = g\mathbf{H}_{0,\mathbb{R}}g^{-1}$ for some $g \in \mathbf{G}(\mathbb{R})^+$ and, for each component X_y^+ of X_y, $(\mathbf{H}_0, g^{-1}X_y^+)$ is a Shimura subdatum component. By Lemma 5.2, there is only one Shimura subdatum component with group $\mathbf{H}_0$. We denote this component by X_0^+. Therefore, $g^{-1}X_y^+ = X_0^+$ for every component X_y^+ of X_y. Hence, X_y is connected and $(\mathbf{H}_y, X_y)$ is a Shimura subdatum component of $(\mathbf{G}, X^+)$.

This establishes property (1) of section 1.D. To establish property (2) of section 1.D, we use the method of [DO22, Prop. 6.3]. All we have to do is understand how fundamental sets in $\mathbf{H}_y$ vary through the $\mathbf{G}(\mathbb{R})$-conjugacy class. A quantitative description of these fundamental sets is given by [DO22, Th. 1.2], but it requires as input a suitable representation ρ and bounds on the lengths of certain vectors in ρ. This input is given in Propositions 6.1 and 7.1, which together generalise [DO22, Prop. 5.1] (which is the case $d = 2$, $e = m = 1$). The representation is constructed in Proposition 6.1, and the construction of vectors w_u with bounds for their lengths is found in Proposition 7.1.

To explain how Theorem 1.2 is used, we outline the proof of Proposition 7.1. Let $L = \mathbb{Z}^{2g}$ and let $V = L_{\mathbb{Q}} = \mathbb{Q}^{2g}$, with the standard action of $\mathbf{G} = \mathbf{GSp}_{2g}$ on V. Choosing a lift $\widetilde{s} \in \pi^{-1}(s)$ induces an isomorphism $L \cong H_1(A_s, \mathbb{Z})$, hence an action of $\mathrm{End}(A_s)$ on L. The polarisation induces a $(D, \dagger)$-skew-Hermitian form ψ on V,

where $D = \mathrm{End}(A_s) \otimes \mathbb{Q}$ and $\dagger$ is the Rosati involution. We use Theorem 1.2 to choose a weakly unitary or weakly symplectic D-basis $\{v_i\}$ for (V, ψ) contained in L. Suitable multiples of the $\{v_i\}$ yield a symplectic or α-unitary $D_{\mathbb{R}}$-basis for $(V_{\mathbb{R}}, \psi)$ (see section 3.B for definitions). The choice of a symplectic or α-unitary $D_{\mathbb{R}}$-basis for $(V_{\mathbb{R}}, \psi)$ is equivalent to the choice of an element $u' \in \mathbf{Sp}_{2g}(\mathbb{R})$ such that $\widetilde{s} \in u' X_0^+$. This element u' is called $\theta^{-1} = uh$ in section 7, and its construction is detailed in Lemmas 7.3 and 7.4.

We then use the bound from Theorem 1.2(iv), via Lemma 7.3(iii), together with the fact that $v_i \in L$, to obtain $\gamma \in \mathbf{Sp}_{2g}(\mathbb{Z})$ such that the entries of the matrices γg and $(\gamma g)^{-1}$ are polynomially bounded (Lemmas 7.5 to 7.7). Since π is $\mathbf{Sp}_{2g}(\mathbb{Z})$-invariant, we still have $\pi^{-1}(s) \cap \gamma g X_0^+ \neq \varnothing$. From γg, we construct a vector (denoted w_u in section 7) suitable for use as input to [DO22, Th. 1.2], which gives the height bound for $y = \rho(b^{-1}u)w_u$.

1.F. Remark on effectivity. — We note that Theorem 1.2 and Theorem 8.5 are effective. As such, the obstructions to effectivity in Theorem 1.5 are (1) its dependence on the (ineffective) Habegger–Pila–Wilkie theorem (as stated in [DR18, Th. 9.1]) from o-minimality and (2) the ineffectivity in [DO22], as explained in Remark 4.3 therein. Obstruction (1) was recently overcome for the André–Oort conjecture for non-compact curves in Hilbert modular varieties by Binyamini and Masser [BM21] using so-called Q-functions. It seems plausible that these techniques could also apply to our setting.

1.G. Outline of the paper. — The paper is in two parts. The first part, sections 2 to 4, proves Theorem 1.2. It deals only with modules over division algebras and skew-Hermitian forms, with no mention of Shimura varieties. The second part, sections 5 to 8, proves Theorem 1.3. The main new ingredient is Theorem 1.2.

In section 2, we introduce terminology around division algebras and their orders, as well as various lemmas used throughout the calculations in sections 3 and 4. In section 3, we define the notion of a skew-Hermitian form on a module over a division algebra with involution and define several notions of well-behaved bases with respect to a skew-Hermitian form. Section 4 consists of the proof of Theorem 1.2, which involves substantial calculations.

Section 5 introduces Shimura data and establishes the basic properties of special subvarieties of simple PEL type I and II in $\mathcal{A}_g$. The representation and vectors required as input for [DO22, Th. 1.2] are constructed in sections 6 and 7, as sketched in section 1.E. The application of Theorem 1.2 is found in section 7, specifically Lemma 7.3. Finally section 8 states some slightly stronger versions of Theorems 1.3 and 1.5 and completes their proofs.

1.H. Notation. — We shall use the following notation for matrices. If A and B are square matrices, we will denote by $A \oplus B$ the block diagonal matrix with blocks A (top-left) and B (bottom-right). We will write $A^{\oplus d}$ to denote the block diagonal matrix $A \oplus \cdots \oplus A$ with A appearing d times.

We shall write $J_2 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and $J_n = J_2^{\oplus n/2}$ for each even positive integer n.

Acknowledgements. — The authors are grateful to the referee for their careful reading of the paper and helpful suggestions and corrections. They are also grateful to Bijay Bhatta, who found several errors in a preprint version of the paper.

2. Division algebras

In this section, we introduce the notation and terminology we shall use for division algebras. A key definition is a norm $|\cdot|_D$ on an $\mathbb{R}$-algebra with positive involution. We establish useful properties of this norm and of the discriminants of orders in division algebras. We also include some broader preliminary lemmas, on discriminants of bilinear forms and versions of Minkowski's second theorem.

In this paper, our main interest will be in division $\mathbb{Q}$-algebras with positive involution of Albert types I and II. However, we have stated many of the definitions and results in this section in greater generality, such as for semisimple algebras over any subfield of $\mathbb{R}$. We do this not only because this greater generality is often natural, but also it is sometimes necessary as we wish to apply the results to $D \otimes_{\mathbb{Q}} \mathbb{R}$ where D is a division $\mathbb{Q}$-algebra, but $D \otimes_{\mathbb{Q}} \mathbb{R}$ might not be a division algebra. We have not stated all results at their greatest possible generality, if doing so would require additional complications while not being required for our application.

Throughout this section, k denotes a subfield of $\mathbb{R}$. Later in the paper, we will usually use $k = \mathbb{Q}$ or $k = \mathbb{R}$. Whenever we say *k-algebra*, we mean a k-algebra of finite dimension. If V is a k-vector space or k-algebra, then $V_{\mathbb{R}}$ denotes $V \otimes_k \mathbb{R}$.

2.A. Semisimple algebras, norms and traces.

As a reference on semisimple algebras, reduced norm and trace, see [Rei75, §9].

Let D be a semisimple k-algebra. Then $D \cong \prod_{i=1}^{s} D_i$ for some simple k-algebras $D_1, \dots, D_s$. For each i, let F_i be the centre of D_i, which is a field.

We write Trd_{D_i/F_i} and Nrd_{D_i/F_i} for the reduced trace and reduced norm respectively of the central simple algebra D_i/F_i. Letting $\mathrm{Tr}_{F_i/k}$ and $\mathrm{Nm}_{F_i/k}$ denote the trace and norm of finite extensions of fields, we define

$$\mathrm{Trd}_{D/k} = \sum_{i=1}^{s} \mathrm{Tr}_{F_i/k} \circ \mathrm{Trd}_{D_i/F_i}, \quad \mathrm{Nrd}_{D/k} = \prod_{i=1}^{s} \mathrm{Nm}_{F_i/k} \circ \mathrm{Nrd}_{D_i/F_i}\,.$$

Note that $\mathrm{Trd}_{D/k}$ and $\mathrm{Nrd}_{D/k}$ are compatible with extension of scalars. By this, we mean that, if K is a field containing k and $D_K = D \otimes_k K$, then $\mathrm{Trd}_{D/k} = \mathrm{Trd}_{D_K/K}|_D$ and similarly for $\mathrm{Nrd}_{D/k}$. This is true even though the simple factors of D might not remain simple after extension of scalars. Note also that $\mathrm{Trd}_{D/k}(ab) = \mathrm{Trd}_{D/k}(ba)$ for all $a, b \in D$.

Suppose that D is a simple k-algebra and let F be the centre of D. Let

$$d = \sqrt{\dim_F(D)} = \mathrm{Trd}_{D/F}(1), \qquad e = [F : k].$$

Then $\dim_k(D) = d^2 e$. We will use the notation F, d, e from this paragraph throughout the paper whenever we talk about simple algebras, without further comment. Note that $\mathrm{Tr}_{D/F}(a) = d\,\mathrm{Trd}_{D/F}(a)$ and $\mathrm{Nm}_{D/F}(a) = \mathrm{Nrd}_{D/F}(a)^d$ for all $a \in D$, where $\mathrm{Tr}_{D/F}$ and $\mathrm{Nm}_{D/F}$ are the non-reduced trace and norm.

2.B. Division algebras with positive involution. — Let D be a semisimple k-algebra. An *involution* $\dagger$ of D means a k-linear map $D \to D$ such that $\dagger \circ \dagger = \mathrm{id}_D$ and $(ab)^\dagger = b^\dagger a^\dagger$ for all $a, b \in D$. (We follow the convention of [Mil05, §8] by requiring involutions of k-algebras to be k-linear. This is important for Lemma 3.1. Thus, with our definition, an "involution of the second kind" of a central simple F-algebra is not an F-algebra involution. However, an involution of the second kind can still be handled within our framework by taking k to be the fixed subfield of F.)

For every $a \in D$, we have $\mathrm{Trd}_{D/k}(a^\dagger) = \mathrm{Trd}_{D/k}(a)$. Consequently, the bilinear form $D \times D \to k$ given by $(a, b) \mapsto \mathrm{Trd}_{D/k}(ab^\dagger)$ is symmetric. The involution $\dagger$ is said to be *positive* if this bilinear form is positive definite (equivalently, if the non-reduced trace bilinear form $(a, b) \mapsto \mathrm{Tr}_{D/k}(ab^\dagger)$ is positive definite).

Division $\mathbb{Q}$-algebras with positive involution $(D, \dagger)$ were classified by Albert into four types, depending on the isomorphism type of $D_\mathbb{R}$ [Mum74, §21, Th. 2].

Type I. — $D = F$, a totally real number field. The involution is trivial. (In this case $D_\mathbb{R} \cong \mathbb{R}^e$.)

Type II. — D is a non-split totally indefinite quaternion algebra over a totally real number field F. (Totally indefinite means that $D_\mathbb{R} \cong \mathrm{M}_2(\mathbb{R})^e$.) The involution is of orthogonal type, meaning that after extending scalars to $\mathbb{R}$ it becomes matrix transpose on each copy of $\mathrm{M}_2(\mathbb{R})$.

Type III. — D is a totally definite quaternion algebra over a totally real number field F. (Totally definite means that $D_\mathbb{R} \cong \mathbb{H}^e$ where $\mathbb{H}$ is Hamilton's quaternions.) The involution is the "canonical involution" defined by $a \mapsto \mathrm{Trd}_{D/F}(a) - a$.

Type IV. — D is a division algebra whose centre is a CM field F. The involution restricts to complex conjugation on F. (In this case $D_\mathbb{R} \cong \mathrm{M}_d(\mathbb{C})^e$.)

2.C. The norm $|\cdot|_D$. — Let $(D, \dagger)$ be a semisimple k-algebra with a positive involution. We define a norm $|\cdot|_D$ on $D_\mathbb{R}$ by:

$$|a|_D = \sqrt{\mathrm{Trd}_{D_\mathbb{R}/\mathbb{R}}(aa^\dagger)}.$$

This is a norm in the sense of a real vector space norm (that is, a length function). Note that $|a^\dagger|_D = |a|_D$ for all $a \in D_\mathbb{R}$.

The norm $|\cdot|_D$ is induced by the inner product $(a, b) \mapsto \mathrm{Trd}_{D_\mathbb{R}/\mathbb{R}}(ab^\dagger)$ on $D_\mathbb{R}$. This inner product (together with an orientation of $D_\mathbb{R}$) also induces a volume form. Whenever we refer to the covolume of a lattice in $D_\mathbb{R}$, we use this volume form. (Note that the covolume is the absolute value of the integral of the volume form over a fundamental domain, so it is independent of the choice of orientation.)

If D is a semisimple k-algebra, then

$$D_\mathbb{R} \cong \prod_{i=1}^{r} \mathrm{M}_{s_i}(\mathbb{K}_i),$$

where $\mathbb{K}_i = \mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$. If D is equipped with a positive involution $\dagger$, then we can choose the isomorphism so that $\dagger$ corresponds to conjugate-transpose on each simple factor [Voi21, Prop. 8.4.7]. Throughout the paper, whenever we choose an isomorphism between $D_{\mathbb{R}}$ and a product of matrix algebras, we implicitly assume that it has this property.

Let $|\cdot|_F$ denote the *Frobenius norm* on any matrix algebra over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$:

$$|M|_F^2 = \sum_{j,k=1}^{s} M_{jk}\overline{M_{jk}}.$$

Then, for any $a = (a_1, \dots, a_r) \in \prod_i \mathrm{M}_{s_i}(\mathbb{K}_i)$, we have

$$|a|_D^2 = \sum_{i=1}^{r} |a_i|_F^2.$$

The following lemma will be used repeatedly throughout sections 3 and 4. It is well-known in the case $D_{\mathbb{R}} = \mathrm{M}_n(\mathbb{R})$ or $\mathrm{M}_n(\mathbb{C})$ – see, for example, [HJ85, p. 291].

Lemma 2.1. — *Let $(D, \dagger)$ be a semisimple k-algebra with positive involution. Then $|ab|_D \leqslant |a|_D |b|_D$ for all $a, b \in D_{\mathbb{R}}$.*

Proof. — Identify $D_{\mathbb{R}}$ with $\prod_{i=1}^{r} \mathrm{M}_{s_i}(\mathbb{K}_i)$ and write

$$a = (a_1, \dots, a_r),\ b = (b_1, \dots, b_r) \in \prod_{i=1}^{r} \mathrm{M}_{s_i}(\mathbb{K}_i).$$

Then

$$|ab|_D^2 = \sum_{i=1}^{r} \|a_i b_i\|_F^2 \leqslant \sum_{i=1}^{r} \|a_i\|_F^2 \|b_i\|_F^2 \leqslant \Big(\sum_{i=1}^{r} \|a_i\|_F\Big)^2 \Big(\sum_{i=1}^{r} \|b_i\|_F\Big)^2 = |a|_D^2 |b|_D^2.$$

This calculation uses the submultiplicativity of the Frobenius norm and the following inequality, valid for all non-negative real numbers $x_1, \dots, x_r, y_1, \dots, y_r$:

$$\sum_{i=1}^{r} x_i y_i \leqslant \Big(\sum_{i=1}^{r} x_i\Big)\Big(\sum_{i=1}^{r} y_i\Big). \tag{2.1}$$

Since the Frobenius norm is less well-known over $\mathbb{H}$, we remark that, just as in the real and complex cases, submultiplicativity of the Frobenius norm follows from the Cauchy–Schwarz inequality

$$\Big(\sum_{j=1}^{s} x_j \overline{y}_j\Big)\Big(\sum_{j=1}^{s} y_j \overline{x}_j\Big) \leqslant \Big(\sum_{j=1}^{s} x_j \overline{x}_j\Big)\Big(\sum_{j=1}^{s} y_j \overline{y}_j\Big) \quad \text{for all } x, y \in \mathbb{K}^n.$$

The Cauchy–Schwarz inequality over $\mathbb{H}$ can be proved by considering the discriminant of the quadratic polynomial $(\sum_{i=1}^{s} x_j t + y_j)(\sum_{i=1}^{n} \overline{x}_j t + \overline{y}_j)$, which is non-negative for all $t \in \mathbb{R}$, and then applying the arithmetic mean–geometric mean inequality to the left hand side. □

We say that a semisimple k-algebra D is $\mathbb{R}$-*split* if $D_{\mathbb{R}} \cong \mathrm{M}_d(\mathbb{R})^e$ for some positive integers d and e. Note that a division $\mathbb{Q}$-algebra with positive involution is $\mathbb{R}$-split if and only if it has type I or II in the Albert classification, and these are the types of algebras that we focus on in this paper.

Lemma 2.2. — *Let $(D,\dagger)$ be an $\mathbb{R}$-split semisimple k-algebra with positive involution and let F be its centre. Then, for all $a \in D_{\mathbb{R}}^{\times}$:*

(i) $|\mathrm{Nrd}_{D_{\mathbb{R}}/F_{\mathbb{R}}}(a)|_D \leqslant d^{(1-d)/2}|a|_D^d$;
(ii) $|\mathrm{Nrd}_{D_{\mathbb{R}}/\mathbb{R}}(a)| \leqslant (de)^{-de/2}|a|_D^{de}$.

Proof. — Identify $D_{\mathbb{R}}$ with $\mathrm{M}_d(\mathbb{R})^e$ and write

$$a = (a_1, \dots, a_e).$$

For each i, the matrix $a_i a_i^t \in \mathrm{M}_d(\mathbb{R})$ is symmetric and positive definite and therefore diagonalisable with positive eigenvalues. Let its eigenvalues be $\lambda_{i1}, \dots, \lambda_{id}$. Note that $|a_i|_F^2 = \mathrm{Tr}(a_i a_i^t) = \lambda_{i1} + \cdots + \lambda_{id}$. By the arithmetic mean–geometric mean inequality,

$$\det(a_i)^{2/d} = \det(a_i a_i^t)^{1/d} = \big(\lambda_{i1} \cdots \lambda_{id}\big)^{1/d} \leqslant d^{-1}\big(\lambda_{i1} + \cdots + \lambda_{id}\big) = d^{-1}|a_i|_F^2. \tag{2.2}$$

(i) We have $\mathrm{Nrd}_{D_{\mathbb{R}}/F_{\mathbb{R}}}(a) = (\det(a_1)I_d, \dots, \det(a_e)I_d)$ where I_d denotes the identity matrix in $\mathrm{M}_d(\mathbb{R})$. Hence

$$\begin{aligned}
|\mathrm{Nrd}_{D_{\mathbb{R}}/F_{\mathbb{R}}}(a)|_D^2 &= \sum_{i=1}^{e} |\det(a_i)I_d|_F^2 = \sum_{i=1}^{e} d|\det(a_i)|^2 \\
&\leqslant \sum_{i=1}^{e} d\big(d^{-1}|a_i|_F^2\big)^d = d^{1-d}\sum_{i=1}^{e}|a_i|_F^{2d} \\
&\leqslant d^{1-d}\Big(\sum_{i=1}^{e}|a_i|^2\Big)^d = d^{1-d}|a|_D^{2d}.
\end{aligned}$$

(ii) Using (2.2) and another application of the AM-GM inequality,

$$|\mathrm{Nrd}_{D_{\mathbb{R}}/\mathbb{R}}(a)|^{2/de} = \Big(\prod_{i=1}^{e}|\det(a_i)|^{2/d}\Big)^{1/e} \leqslant e^{-1}\sum_{i=1}^{e} d^{-1}|a_i|_F^2 = (de)^{-1}|a|_D^2. \qquad \square$$

2.D. The Hermite constant and Minkowski's theorems. — Let γ_n denote the Hermite constant for $\mathbb{R}^n$, that is, the smallest positive real number such that the following holds: *For every lattice L in $\mathbb{R}^n$ with the Euclidean norm and volume form, there exists a vector $v \in L$ satisfying $|v| \leqslant \sqrt{\gamma_n}\,\mathrm{covol}(L)^{1/n}$.*

It is immediate from the definition that $\gamma_n \geqslant 1$ for all n. As a consequence of Minkowski's theorem on convex bodies,

$$\gamma_n \leqslant 4\mathcal{V}_n^{-2/n} = \tfrac{4}{\pi}\,\Gamma(\tfrac{n}{2}+1)^{2/n}, \tag{2.3}$$

where $\mathcal{V}_n$ denotes the volume of the unit ball in $\mathbb{R}^n$.

Lemma 2.3. — *For all positive integers n, we have $\gamma_n \leqslant n$.*

Proof. — According to [AQ97, Th. 1.5], $\Gamma(x) \leqslant x^{x-1}$ for all real numbers $x > 1$. Hence

$$\Gamma(x+1) = x\Gamma(x) < x^x$$

for all $x > 1$. Furthermore, for $x = 1$, we have $\Gamma(x+1) = 1 = x^x$. Thus $\Gamma(x+1) \leqslant x^x$ for all $x \geqslant 1$. Plugging this into (2.3), we obtain $\gamma_n \leqslant (4/\pi) \cdot (n/2) < n$ for all $n \geqslant 2$.

It is clear that $\gamma_1 = 1$ [Cas97, App.], so the lemma is also true for $n = 1$. □

Lemma 2.3 is not optimal for large n. Indeed, our proof itself shows that $\gamma_n \leqslant \frac{2}{\pi} n$ for $n \geqslant 2$. Using Stirling's approximation to the Gamma function, one can obtain $\gamma_n \leqslant 2n/\pi e + o(n)$ as $n \to +\infty$. However we have chosen to use Lemma 2.3 because we need a simple bound for the Hermite constant which is valid for all $n \geqslant 1$, without hidden constants or special cases for small n, as we wish to avoid fiddly special cases when calculating the (non-optimal) constants in Proposition 4.5.

A version of Minkowski's second theorem for the Euclidean norm also holds with the Hermite constant:

Theorem 2.4 ([Cas97, Chap. VIII, Th. 1]). — *For every lattice L in $\mathbb{R}^n$ with the Euclidean norm and volume form, there exist vectors $e_1, \dots, e_n \in L$ which form a basis for $\mathbb{R}^n$ and which satisfy $|e_1| \cdots |e_n| \leqslant \gamma_n^{n/2} \operatorname{covol}(L)$.*

With some book-keeping, we can obtain a version of Theorem 2.4 for vector spaces over a division $\mathbb{Q}$-algebra. This is the same method as the proof of a version of Minkowski's second theorem over number fields in [BG06, C.2.18].

Proposition 2.5. — *Let D be a division $\mathbb{Q}$-algebra. Let V be a left D-vector space of dimension m. Let L be a $\mathbb{Z}$-lattice in V. Let $|\cdot|$ be any norm on $V_{\mathbb{R}}$ induced by an inner product, and use the associated volume form to define* $\operatorname{covol}(L)$.

Then there exists a D-basis $w_1, \dots, w_m$ for V such that:

(i) $w_1, \dots, w_m \in L$;
(ii) $|w_1||w_2| \cdots |w_m| \leqslant \gamma_{[D:\mathbb{Q}]m}^{m/2} \operatorname{covol}(L)^{1/[D:\mathbb{Q}]}$.

Proof. — Let $n = \dim_{\mathbb{Q}}(V) = [D : \mathbb{Q}]m$. Choose $e_1, \dots, e_n \in L$ as in Theorem 2.4. Order the e_i so that $|e_i| \leqslant |e_{i+1}|$ for all $i = 1, \dots, n-1$.

For $i = 1, \dots, m$, let q_i denote the smallest positive integer q such that the D-span of $e_1, \dots, e_q$ has D-dimension equal to i. Let $w_i = e_{q_i}$. By construction, for each i, the D-span of $w_1, \dots, w_i$ has D-dimension equal to i. Hence $w_1, \dots, w_m$ is a D-basis for V.

For $1 \leqslant i \leqslant m$, the vectors $e_1, \dots, e_{q_i - 1}$ are contained in a D-vector space of D-dimension $i-1$, so they are contained in a $\mathbb{Q}$-vector space of $\mathbb{Q}$-dimension at most $[D : \mathbb{Q}](i-1)$. These vectors are $\mathbb{Q}$-linearly independent, so

$$q_i - 1 \leqslant [D : \mathbb{Q}](i-1).$$

Since the lengths $|e_i|$ are increasing, we deduce that

$$|w_i|^{[D:\mathbb{Q}]} \leqslant |e_{[D:\mathbb{Q}](i-1)+1}|^{[D:\mathbb{Q}]} \leqslant \prod_{j=1}^{[D:\mathbb{Q}]} |e_{[D:\mathbb{Q}](i-1)+j}|.$$

Hence by Theorem 2.4,

$$\prod_{i=1}^{m} |w_i|^{[D:\mathbb{Q}]} \leqslant \prod_{i=1}^{n} |e_i| \leqslant \gamma_n^{n/2} \operatorname{covol}(L). \qquad \square$$

Let D be a division $\mathbb{Q}$-algebra, R an order in D and L a torsion-free R-module of rank m. Combining Proposition 2.5 with Theorem 2.4 applied to R and Hadamard's inequality, we could prove that there exist $w_1, \dots, w_m \in L$ forming a D-basis for $D \otimes_R L$ and satisfying $[L : Rw_1 + \cdots + Rw_m] \leqslant C_9 |\operatorname{disc}(R)|^{m/2}$. However this method of proof gives a constant $C_9 > 1$, so this is weaker than Theorem 1.1.

2.E. Discriminants of bilinear forms. — If Λ is a $\mathbb{Z}$-module, we write $\Lambda_{\mathbb{Q}}$ for $\Lambda \otimes_{\mathbb{Z}} \mathbb{Q}$. If Λ is free of finite rank and $\phi\colon \Lambda_{\mathbb{Q}} \times \Lambda_{\mathbb{Q}} \to \mathbb{Q}$ is a bilinear form, we write $\operatorname{disc}(\Lambda, \phi)$ for the determinant of the matrix $(\phi(e_i, e_j))_{i,j}$ where $\{e_1, \dots, e_n\}$ is a $\mathbb{Z}$-basis for Λ (the determinant is independent of the choice of basis).

Lemma 2.6. — *Let L be a free $\mathbb{Z}$-module of finite rank and let $\phi\colon L \times L \to \mathbb{Z}$ be a non-degenerate bilinear form. Let $M \subset L$ be a $\mathbb{Z}$-submodule such that $\phi_{|M \times M}$ is non-degenerate. Let*

$$M^{\perp} = \{x \in L : \phi(x, y) = 0 \text{ for all } y \in M\}.$$

Then

(i) $[L : M + M^{\perp}] \leqslant |\operatorname{disc}(M, \phi)|$*; and*
(ii) $|\operatorname{disc}(M^{\perp}, \phi)| \leqslant |\operatorname{disc}(L, \phi)||\operatorname{disc}(M, \phi)|$.

Proof. — Since $\phi_{|M \times M}$ is non-degenerate, $L_{\mathbb{Q}} = M_{\mathbb{Q}} \oplus M_{\mathbb{Q}}^{\perp}$. Let $\pi\colon L_{\mathbb{Q}} \to M_{\mathbb{Q}}$ denote the projection with kernel $M_{\mathbb{Q}}^{\perp}$.

If $x \in L$ and $\pi(x) \in M$, then $x - \pi(x) \in \ker(\pi) \cap L = M^{\perp}$. Hence $x \in M + M^{\perp}$. Conversely, if $x \in M + M^{\perp}$, it is clear that $\pi(x) \in M$. Thus $\pi^{-1}(M) = M + M^{\perp}$.

Let

$$M^* = \{x \in M_{\mathbb{Q}} : \phi(x, y) \in \mathbb{Z} \text{ for all } y \in M\}.$$

If $x \in L$, then $\phi(\pi(x), y) = \phi(x, y) \in \mathbb{Z}$ for all $y \in M$ so $\pi(x) \in M^*$. Thus $\pi(L) \subset M^*$. Thus we obtain

$$L/(M + M^{\perp}) = L/\pi^{-1}(M) \cong \pi(L)/M \subset M^*/M.$$

It is well-known that $[M^* : M] = |\operatorname{disc}(M, \phi)|$, so this proves (i).

Since M and $M^{\perp}$ are orthogonal with respect to ϕ,

$$\begin{aligned} |\operatorname{disc}(M, \phi)||\operatorname{disc}(M^{\perp}, \phi)| &= |\operatorname{disc}(M + M^{\perp}, \phi)| \\ &= [L : M + M^{\perp}]^2 |\operatorname{disc}(L, \phi)| \leqslant |\operatorname{disc}(M, \phi)|^2 |\operatorname{disc}(L, \phi)|. \end{aligned}$$

Since $|\operatorname{disc}(M, \phi)| \neq 0$, this proves (ii). $\square$

2.F. Orders and discriminants. — Let $k = \mathbb{Q}$ or $\mathbb{R}$. If V is a finite-dimensional k-vector space, then a $\mathbb{Z}$-*lattice* in V means a $\mathbb{Z}$-submodule $L \subset V$ such that the natural map $L \otimes_{\mathbb{Z}} k \to V$ is an isomorphism.

Let D be a semisimple $\mathbb{Q}$-algebra. An *order* in D is a $\mathbb{Z}$-lattice in D which is also a subring. Note that if V is a D-vector space and L is a $\mathbb{Z}$-lattice in V, then $\mathrm{Stab}_D(L) = \{a \in D : aL \subset L\}$ is an order in D. (This is proved on [Rei75, p. 109] when $V = D$, and the proof generalises.)

If R is an order in D, the *discriminant* $\mathrm{disc}(R)$ is defined to be the discriminant of the k-bilinear form $(a, b) \mapsto \mathrm{Tr}_{D/k}(ab)$ on R, where $\mathrm{Tr}_{D/\mathbb{Q}}$ is the *non-reduced* trace. The trace form of a semisimple algebra is non-degenerate, so $\mathrm{disc}(R) \neq 0$. Furthermore, $\mathrm{Tr}_{D/\mathbb{Q}}(a) \in \mathbb{Z}$ for all $a \in R$, so $\mathrm{disc}(R) \in \mathbb{Z}$.

If D is a simple $\mathbb{Q}$-algebra, then $\mathrm{Trd}_{D/\mathbb{Q}}(a) \in \mathbb{Z}$ for all $a \in R$ [Rei75, Th. 10.1]. Since $\mathrm{Tr}_{D/\mathbb{Q}} = d\,\mathrm{Trd}_{D/\mathbb{Q}}$, it follows that $\mathrm{disc}(R) \in d^{d^2e}\mathbb{Z}$ so

$$|\mathrm{disc}(R)| \geqslant d^{d^2e}. \tag{2.4}$$

Now suppose that $(D, \dagger)$ is a simple $\mathbb{Q}$-algebra with a positive involution. According to [DO22, Lem. 5.6], for any order $R \subset D$, $|\mathrm{disc}(R)|$ is equal to the discriminant of the symmetric bilinear form $(a, b) \mapsto \mathrm{Tr}_{D/k}(ab^\dagger)$. Consequently, $|\mathrm{disc}(R)|$ is equal to d^{d^2e} multiplied by the discriminant on R of the positive definite bilinear form which induces the norm $|\cdot|_D$. We conclude that

$$|\mathrm{disc}(R)| = d^{d^2e}\,\mathrm{covol}(R)^2. \tag{2.5}$$

For an order R in a simple $\mathbb{Q}$-algebra D, let R^* denote the dual lattice

$$R^* = \{a \in D : \mathrm{Trd}_{D/\mathbb{Q}}(ab) \in \mathbb{Z} \text{ for all } b \in R\}.$$

Lemma 2.7. — *Let D be a semisimple k-algebra and let R be an order in D. Let F be the centre of D and let $\mathcal{O}$ be an order in F which contains $R \cap F$. Then*

$$[\mathcal{O} : R \cap F]^2\,|\mathrm{disc}(\mathcal{O}R)| \leqslant |\mathrm{disc}(R)|.$$

Proof. — This follows from the facts $\mathcal{O} + R \subset \mathcal{O}R$ and $[\mathcal{O} + R : R] = [\mathcal{O} : R \cap F]$. □

Lemma 2.8. — *Let D be a simple $\mathbb{Q}$-algebra. Let F be the centre of D and let $\mathcal{O}_F$ be the maximal order of F. Let S be an order in D which contains $\mathcal{O}_F$. Define S^* analogously to R^*. Then there exists an ideal $I \subset \mathcal{O}_F$ such that $IS^* \subset S$ and*

$$\mathrm{Nm}(I) \leqslant d^{-d^2e}|\mathrm{disc}(S)|.$$

Proof. — Let $I = \{x \in \mathcal{O}_F : xS^* \subset S\}$, that is, the annihilator of the finite $\mathcal{O}_F$-module S^*/S. By the structure theorem for finitely generated torsion modules over a Dedekind domain, there is an isomorphism of $\mathcal{O}_F$-modules

$$S^*/S \cong \mathcal{O}_F/I_1 \oplus \mathcal{O}_F/I_2 \oplus \cdots \oplus \mathcal{O}_F/I_r$$

for some $\mathcal{O}_F$-ideals $I_1, I_2, \ldots, I_r$. We have $I = I_1 \cap I_2 \cap \cdots \cap I_r \supset I_1 I_2 \cdots I_r$ and so

$$\mathrm{Nm}(I) \leqslant \mathrm{Nm}(I_1)\,\mathrm{Nm}(I_2) \cdots \mathrm{Nm}(I_r) = [S^* : S].$$

The index $[S^* : S]$ is equal to the absolute value of the discriminant of S with respect to the reduced trace form. Thus $[S^* : S] = d^{-d^2e}|\mathrm{disc}(S)|$. □

Lemma 2.9. — *Let D be a simple $\mathbb{Q}$-algebra. Let F be the centre of D and let $\mathcal{O}_F$ be the maximal order of F. Let R be an order in D. Let $S = \mathcal{O}_F R$. Let $\mathfrak{c}$ be the conductor of $R \cap F$ (as an order in the number field F). Then*

$$\mathfrak{c}S \subset R \quad \text{and} \quad \mathfrak{c}R^* \subset S^*.$$

Proof. — From the definitions of S and $\mathfrak{c}$,

$$\mathfrak{c}S = \mathfrak{c}\mathcal{O}_F R \subset (R \cap F)R \subset R.$$

If $c \in \mathfrak{c}$ and $a \in R^*$, then for all $b \in S$ we have

$$\mathrm{Trd}_{D/\mathbb{Q}}((ca)b) = \mathrm{Trd}_{D/\mathbb{Q}}(a(cb)) \in \mathbb{Z}$$

because c is in the centre of D and $cb \in \mathfrak{c}S \subset R$. Thus $ca \in S^*$. □

Lemma 2.10. — *Let D be a division $\mathbb{Q}$-algebra and let V be a left D-vector space of dimension m. Let L be a $\mathbb{Z}$-lattice in V and consider the order $R = \mathrm{Stab}_D(L)$ of D. Let $S = \mathrm{End}_R(L)$ denote the ring of endomorphisms of L commuting with R. Then*

$$|\mathrm{disc}(S)| \leqslant |\mathrm{disc}(R)|^{(d^2em+1)m^2}.$$

Proof. — By Theorem 1.1, there is a D-basis $v_1, \dots, v_m$ for V such that $v_1, \dots, v_m \in L$ and

$$[L : Rv_1 + \cdots + Rv_m] \leqslant |\mathrm{disc}(R)|^{m/2}. \tag{2.6}$$

Let $N = [L : Rv_1 + \cdots + Rv_m]$ and $s = \dim_{\mathbb{Q}}(\mathrm{End}_D(V)) = d^2em^2$.

Using the D-basis $v_1, \dots, v_m$, we identify $\mathrm{End}_D(V)$ with $\mathrm{M}_m(D^{\mathrm{op}})$. Note that $\mathrm{End}_R(L)$ and $\mathrm{M}_m(R^{\mathrm{op}})$ are both $\mathbb{Z}$-lattices in $\mathrm{End}_D(V)$.

For every $a \in \mathrm{M}_m(R^{\mathrm{op}}) \subset \mathrm{End}_D(V)$, we have

$$aNL \subset a(Rv_1 + \cdots + Rv_m) \subset Rv_1 + \cdots + Rv_m \subset L.$$

Hence $Na \in \mathrm{End}_R(L)$.

Thus $N\mathrm{M}_m(R^{\mathrm{op}}) \subset \mathrm{End}_R(L)$. Therefore

$$|\mathrm{disc}(S)| \leqslant N^{2s}|\mathrm{disc}(\mathrm{M}_m(R^{\mathrm{op}}))| = N^{2s}|\mathrm{disc}(R)|^{m^2}.$$

Combining this with the bound for N from (2.6) proves the lemma. □

2.G. Anti-symmetric elements in division algebras of type II. — If $(D, \dagger)$ is a division $\mathbb{Q}$-algebra with involution, we define

$$D^- = \{a \in D : a^\dagger = -a\}.$$

If $\psi\colon V \times V \to D$ is a $(D, \dagger)$-skew-Hermitian form on a D-vector space V and $x \in V$, then $\psi(x, x) \in D^-$, so D^- is important for the study of weakly unitary bases (see section 3.A for the definition of $(D, \dagger)$-skew-Hermitian forms).

Let $(D,\dagger)$ be a division $\mathbb{Q}$-algebra with a positive involution of Albert type II. Choose an isomorphism $D_{\mathbb{R}} \cong \mathrm{M}_2(\mathbb{R})^e$ (as always, we implicitly assume that $\dagger$ corresponds to matrix transpose on each factor). Then $D_{\mathbb{R}}^-$ consists of those elements of $\mathrm{M}_2(\mathbb{R})^e$ in which all matrices are anti-symmetric. Hence $D_{\mathbb{R}}^-$ is a free $F_{\mathbb{R}}$-module of rank 1, so D^- is a 1-dimensional F-vector space. The following lemma can be proved by calculations in $D_{\mathbb{R}} \cong \mathrm{M}_2(\mathbb{R})^e$.

Lemma 2.11. — *Let $(D,\dagger)$ be a division $\mathbb{Q}$-algebra with a positive involution of type II. Let F be the centre of D.*

(i) *If $a,b \in D^-$, then $ab \in F$.*
(ii) *If $a \in D$ and $b \in D^-$, then $aba^\dagger = \mathrm{Nrd}_{D/F}(a)b$.*

Lemma 2.12. — *Let $(D,\dagger)$ be a division $\mathbb{Q}$-algebra with a positive involution of type II. Let R be an order in D and let $\eta \in \mathbb{Z}_{>0}$ be a positive integer such that $\eta R^\dagger \subset R$. Then there exists $\omega \in D$ such that:*

(i) $\omega \in D^- \smallsetminus \{0\}$;
(ii) $\omega R^* \subset R$ *and* $R^*\omega \subset R$;
(iii) $|\omega|_D \leqslant 2^{-4}\gamma_e^{1/2}\eta^7 |\mathrm{disc}(R)|^{2/e}$.

Proof. — Let F be the centre of D and let $\mathcal{O}_F$ be the maximal order of F. Let $\mathfrak{c} = \{\alpha \in \mathcal{O}_F : \alpha\mathcal{O}_F \subset R \cap F\}$ be the conductor of the order $R \cap F$ in $\mathcal{O}_F$. By [DCD00, (2)], we have the following inclusion of ideals in $\mathbb{Z}$:

$$\mathrm{disc}_{F/\mathbb{Q}}(R \cap F) \subseteq \mathrm{Nm}_{F/\mathbb{Q}}(\mathfrak{c})\,\mathrm{disc}_{F/\mathbb{Q}}(\mathcal{O}_F).$$

This leads to the following inequality of integers:

$$\mathrm{Nm}(\mathfrak{c})|\mathrm{disc}(\mathcal{O}_F)| \leqslant |\mathrm{disc}(R \cap F)|.$$

Since also $|\mathrm{disc}(R \cap F)| = [\mathcal{O}_F : R \cap F]^2 |\mathrm{disc}(\mathcal{O}_F)|$, we deduce that

$$\mathrm{Nm}(\mathfrak{c}) \leqslant [\mathcal{O}_F : R \cap F]^2.$$

Let $S = \mathcal{O}_F R$ and $S^- = S \cap D^-$. Let I be the ideal of $\mathcal{O}_F$ given by Lemma 2.8 applied to S. Let $J = \mathfrak{c}^2 I$ (as a product of ideals of $\mathcal{O}_F$). Then by Lemma 2.9,

$$JSR^* = \mathfrak{c}SI\mathfrak{c}R^* \subset \mathfrak{c}SIS^* \subset \mathfrak{c}SS \subset \mathfrak{c}S \subset R,$$
$$R^*JS = \mathfrak{c}I\mathfrak{c}R^*S \subset \mathfrak{c}IS^*S \subset \mathfrak{c}SS \subset \mathfrak{c}S \subset R.$$

Hence if we choose $\omega \in JS \cap D^- \smallsetminus \{0\} = JS^- \smallsetminus \{0\}$, then it will satisfy (i) and (ii).

Since S^- is a non-zero $\mathcal{O}_F$-submodule of an F-vector space of dimension 1, we can write $S^- = I^-\alpha$ for some ideal $I^- \subset \mathcal{O}_F$ and some $\alpha \in D^-$, then use the multiplicativity of ideal norms in $\mathcal{O}_F$ to conclude that

$$\mathrm{covol}(JS^-) = \mathrm{Nm}(J)\,\mathrm{covol}(S^-),$$

where we measure covolumes in $D_{\mathbb{R}}^-$ by the volume form associated with the restriction of the inner product $\mathrm{Trd}_{D_{\mathbb{R}}/\mathbb{R}}(ab^\dagger)$.

Let $S^+ = \{a \in S : a^\dagger = a\}$. Then $S^+ \cap S^- = \{0\}$. Thus the sum $S^+ + S^-$ is direct. This sum is also orthogonal because, if $a \in S^+$ and $b \in S^-$, then

$$\operatorname{Trd}_{D/\mathbb{Q}}(ab^\dagger) = \operatorname{Trd}_{D/\mathbb{Q}}((ab^\dagger)^\dagger) = \operatorname{Trd}_{D/\mathbb{Q}}(ba^\dagger) = -\operatorname{Trd}_{D/\mathbb{Q}}(ab^\dagger)$$

so $\operatorname{Trd}_{D/\mathbb{Q}}(ab^\dagger) = 0$.

For every $a \in S$, we have $\eta a^\dagger \in \eta(\mathcal{O}_F R)^\dagger = \mathcal{O}_F \eta R^\dagger \subset \mathcal{O}_F R = S$. Hence

$$2\eta a = (\eta a + \eta a^\dagger) + (\eta a - \eta a^\dagger) \in S^+ + S^-.$$

Thus $2\eta S \subset S^+ \oplus S^-$, so

$$\operatorname{covol}(S^+)\operatorname{covol}(S^-) = \operatorname{covol}(S^+ \oplus S^-) \leqslant \operatorname{covol}(2\eta S) = 2^{4e}\eta^{4e}\operatorname{covol}(S). \tag{2.7}$$

Here we measure covolumes in both $D_{\mathbb{R}}^-$ and $S^+ \otimes_{\mathbb{Z}} \mathbb{R}$ by the volume forms associated with the restriction of the inner product $\operatorname{Trd}_{D_{\mathbb{R}}/\mathbb{R}}(ab^\dagger)$.

For all $a, b \in S$, $\eta ab^\dagger \in S$ and so $\operatorname{Trd}_{D/\mathbb{Q}}(ab^\dagger) \in \eta^{-1}\mathbb{Z}$. Consequently, $\operatorname{covol}(S^+) \geqslant \eta^{-\operatorname{rk}_{\mathbb{Z}}(S^+)} = \eta^{-3e}$ so by (2.5) applied to S and (2.7),

$$\operatorname{covol}(S^-) \leqslant \eta^{3e} \cdot 2^{4e}\eta^{4e}\operatorname{covol}(S) = 2^{4e}\eta^{7e} \cdot 2^{-2e}|\operatorname{disc}(S)|^{1/2}.$$

Therefore, using Lemma 2.8,

$$\begin{aligned}\operatorname{covol}(JS^-) &= \operatorname{Nm}(\mathfrak{c})^2 \operatorname{Nm}(I)\operatorname{covol}(S^-) \\ &\leqslant [\mathcal{O}_F : R \cap F]^4 \cdot 2^{-4e}|\operatorname{disc}(S)| \cdot 2^{2e}\eta^{7e}|\operatorname{disc}(S)|^{1/2} \\ &= 2^{-2e}\eta^{7e}[\mathcal{O}_F : R \cap F]^4|\operatorname{disc}(S)|^{3/2}.\end{aligned}$$

Applying (2.4) to S, we see that $|\operatorname{disc}(S)| \geqslant 2^{4e}$. Using Lemma 2.7, we deduce that

$$\operatorname{covol}(JS^-) \leqslant 2^{-2e}|\operatorname{disc}(S)|^{-1/2}\eta^{7e}[\mathcal{O}_F : R \cap F]^4|\operatorname{disc}(S)|^2 = 2^{-4e}\eta^{7e}|\operatorname{disc}(R)|^2.$$

Since JS^- is a free $\mathbb{Z}$-module of rank e, there exists $\omega \in JS^- \smallsetminus \{0\}$ with

$$|\omega|_D \leqslant \sqrt{\gamma_e}\operatorname{covol}(JS^-)^{1/e} \leqslant \sqrt{\gamma_e} \cdot 2^{-4}\eta^7|\operatorname{disc}(R)|^{2/e}. \qquad \square$$

3. Skew-Hermitian forms over division algebras

In this section, we introduce the notion of a $(D, \dagger)$-skew-Hermitian form on a vector space over a division algebra D with an involution, and explain how this is related to skew-symmetric forms over the base field. We define several notions of good behaviour for bases relative to $(D, \dagger)$-skew-Hermitian forms, such as symplectic and unitary bases and a weakened version of these notions. Finally we prove the existence of norms on D-vector spaces, which we call D-norms, which behave well relative to the action of D and to a $(D, \dagger)$-skew-Hermitian form.

As in section 2, we are interested in applying the results of this section when $(D, \dagger)$ is either a division $\mathbb{Q}$-algebra with a positive involution of type I or II, or the semisimple $\mathbb{R}$-algebra which arises from such a $\mathbb{Q}$-algebra by extending scalars to $\mathbb{R}$, but we state the results in greater generality whenever it is convenient.

3.A. Skew-Hermitian forms. — Let k be any field. Let $(D,\dagger)$ be a semisimple k-algebra with an involution. Let V be a left D-module.

A $(D,\dagger)$-*skew-Hermitian form* on V is a k-bilinear map $\psi\colon V\times V\to D$ which satisfies

$$\psi(y,x)=-\psi(x,y)^\dagger \text{ and } \psi(ax,by)=a\psi(x,y)b^\dagger$$

for all $a,b\in D$ and $x,y\in V$. We say that a $(D,\dagger)$-skew-Hermitian form ψ is *non-degenerate* if, for every $x\in V\smallsetminus\{0\}$, there exists $y\in V$ such that $\psi(x,y)\neq 0$.

A $(D,\dagger)$-*compatible skew-symmetric form* on V is a skew-symmetric k-bilinear map $\phi\colon V\times V\to k$ which satisfies

$$\phi(ax,y)=\phi(x,a^\dagger y)$$

for all $a\in D$ and $x,y\in V$. A pair (V,ϕ), where ϕ is a $(D,\dagger)$-compatible skew-symmetric form, is called a symplectic $(D,\dagger)$-module in [Mil05, §8].

Lemma 3.1. — *Let $(D,\dagger)$ be a semisimple k-algebra with an involution. Let V be a left D-module. Then the map $\psi\mapsto \mathrm{Trd}_{D/k}\circ\psi$ is a bijection between the set of $(D,\dagger)$-skew-Hermitian forms on V and the set of $(D,\dagger)$-compatible skew-symmetric forms on V.*

Proof. — It is clear that, if ψ is a $(D,\dagger)$-skew-Hermitian form on V, then $\mathrm{Trd}_{D/k}\psi$ is a $(D,\dagger)$-compatible skew-symmetric form.

Let ϕ be a $(D,\dagger)$-compatible skew-symmetric form. We shall show that there is a unique $(D,\dagger)$-skew-Hermitian form on V such that $\phi=\mathrm{Trd}_{D/k}\psi$.

For each $x,y\in V$, define a k-linear map $\alpha_{x,y}\colon D\to k$ by $\alpha_{x,y}(a)=\phi(ax,y)$. Because D is a semisimple k-algebra, $(a,b)\mapsto \mathrm{Trd}_{D/k}(ab)$ is a non-degenerate bilinear form $D\times D\to k$ [Rei75, Th. 9.26]. Hence there exists a unique element $\beta_{x,y}\in D$ such that

$$\alpha_{x,y}(a)=\mathrm{Trd}_{D/k}(a\beta_{x,y}) \text{ for all } a\in D.$$

Define $\psi(x,y)=\beta_{x,y}$. Using the uniqueness of the elements $\beta_{x,y}$, it is clear that the resulting function $\psi\colon V\times V\to D$ is k-bilinear.

If $a,b\in D$ and $x,y\in V$, then

$$\mathrm{Trd}_{D/k}(ab\beta_{x,y})=\alpha_{x,y}(ab)=\phi(abx,y)=\alpha_{bx,y}(a)=\mathrm{Trd}_{D/k}(a\beta_{bx,y}).$$

By uniqueness of $\beta_{bx,y}$, we deduce that ψ is D-linear in the first variable.

If $a\in D$ and $x,y\in V$, then

$$\mathrm{Trd}_{D/k}(a\beta_{x,y})=\phi(ax,y)=-\phi(a^\dagger y,x)=-\mathrm{Trd}_{D/k}(a^\dagger\beta_{y,x})=-\mathrm{Trd}_{D/k}(a\beta_{y,x}^\dagger).$$

Again by uniqueness of $\beta_{bx,y}$, $\psi(x,y)=-\psi(y,x)^\dagger$.

Since ψ is D-linear in the first variable and satisfies $\psi(x,y)=-\psi(y,x)^\dagger$, it is also $(D,\dagger)$-anti-linear in the second variable. Thus it is $(D,\dagger)$-skew-Hermitian. □

Lemma 3.2. — *Let $(D,\dagger)$ be a semisimple k-algebra with an involution. Let V be a left D-module. Let $\psi\colon V\times V\to k$ be a $(D,\dagger)$-skew-Hermitian form and let $\phi=\mathrm{Trd}_{D/k}\psi\colon V\times V\to k$.*

Let $W \subset V$ be a left D-submodule and define

$$W_\psi^\perp = \{x \in V : \psi(w,x) = 0 \text{ for all } w \in W\},$$
$$W_\phi^\perp = \{x \in V : \phi(w,x) = 0 \text{ for all } w \in W\}.$$

Then $W_\psi^\perp = W_\phi^\perp$.

In particular, $W_\phi^\perp$ is a left D-submodule of V.

Proof. — It is clear that $W_\psi^\perp \subset W_\phi^\perp$.

If $x \in W_\phi^\perp$ and $w \in W$ then, for all $a \in D$, we have $aw \in W$ and so

$$\mathrm{Trd}_{D/\mathbb{Q}}(a\psi(w,x)) = \mathrm{Trd}_{D/\mathbb{Q}}(\psi(aw,x)) = \phi(aw,x) = 0.$$

By the non-degeneracy of the reduced trace form, it follows that $\psi(w,x) = 0$, that is, $x \in W_\psi^\perp$. Thus $W_\phi^\perp \subset W_\psi^\perp$. □

Corollary 3.3. — *Let $(D,\dagger)$ be a semisimple k-algebra with an involution. Let V be a left D-module. Let $\psi\colon V \times V \to k$ be a $(D,\dagger)$-skew-Hermitian form and let $\phi = \mathrm{Trd}_{D/k}\psi\colon V \times V \to k$. Then ψ is non-degenerate if and only if ϕ is non-degenerate.*

Proof. — Apply Lemma 3.2 to $W = V$. □

3.B. Weakly symplectic and weakly unitary bases. — Let k be a field satisfying $\mathrm{char}(k) \neq 2$ and let $(D,\dagger)$ be a semisimple k-algebra with an involution. Let V be a free left D-module and let $\psi\colon V \times V \to D$ be a $(D,\dagger)$-skew-Hermitian form.

We will now define special properties relative to ψ which may be possessed by a basis of V. The notion of (weakly) symplectic basis is useful when D a division $\mathbb{Q}$-algebra of type I or k^e, and the notion of (weakly) unitary basis is useful when D is a division $\mathbb{Q}$-algebra of type II or $\mathrm{M}_2(k)^e$.

We say that a D-basis $v_1,\dots,v_m$ for V is *weakly symplectic* if $\psi(v_i,v_j) = 0$ for all i,j except when $\{i,j\} = \{2k-1,2k\}$ for some $k \in \mathbb{Z}$. If ψ is non-degenerate, then this implies that $\psi(v_{2k-1},v_{2k}) \neq 0$ for all k.

We say that a D-basis $v_1,\dots,v_m$ is *symplectic* if ψ is non-degenerate, the basis is weakly symplectic and furthermore, $\psi(v_{2k-1},v_{2k}) = 1$ for all k. When D is a field and $\dagger = \mathrm{id}$, a $(D,\dagger)$-skew-Hermitian form is the same thing as a symplectic form and this definition agrees with the usual definition of symplectic basis.

We say that a D-basis $v_1,\dots,v_m$ is *weakly unitary* if $\psi(v_i,v_j) = 0$ for all $i,j \in \{1,\dots,m\}$ such that $i \neq j$. If ψ is non-degenerate, then this implies that $\psi(v_i,v_i) \neq 0$ for all i.

For a general division algebra with involution $(D,\dagger)$, there is no canonical choice of a non-zero element of D^-, so there is no natural definition of "unitary basis" with respect to a $(D,\dagger)$-skew-Hermitian form. In the special case $D_0 = \mathrm{M}_d(k)^e$ with d even, let us define

$$\omega_0 = (J_d,\dots,J_d) \in D_0^-,$$

where $J_d \in \mathrm{M}_d(k)$ was defined in section 1.H. If V is a free left D_0-module equipped with a (D_0, t)-skew-Hermitian form ψ_0, then we say that a left D_0-basis $v_1, \dots, v_m$ of V is *unitary* if it is weakly unitary and $\psi(v_i, v_i) = \omega_0$ for all $i = 1, \dots, m$.

If $(D, \dagger)$ is a division $\mathbb{Q}$-algebra with positive involution of type II, $\alpha\colon (D_{0,\mathbb{R}}, \dagger) \to (D_{\mathbb{R}}, t)$ is an isomorphism of $\mathbb{R}$-algebras with involution, and V is a left D-vector space equipped with a $(D, \dagger)$-skew-Hermitian form ψ, then we say that a left $D_{\mathbb{R}}$-basis for $V_{\mathbb{R}}$ is *α-unitary* if it forms a unitary $D_{0,\mathbb{R}}$-basis for $V_{\mathbb{R}}$ viewed as a $D_{0,\mathbb{R}}$-module via α, with respect to the $(D_{0,\mathbb{R}}, t)$-skew-Hermitian form $\alpha^{-1} \circ \psi\colon V_{\mathbb{R}} \times V_{\mathbb{R}} \to D_{0,\mathbb{R}}$. The elements v_i of an α-unitary basis satisfy $\psi(v_i, v_i) = \alpha(\omega_0)$.

As an aside, which will be used in later calculations, we remark that, for any $a \in D_0$, the entries of the matrices which make up $a\omega_0$ are (up to signs) a permutation of the matrix entries making up a. Hence

$$|a\omega_0|_{D_0} = |a|_{D_0}. \tag{3.1}$$

The following lemma shows how we can adjust a weakly symplectic or weakly unitary basis to become symplectic or α-unitary. Note that it works only over $D_{\mathbb{R}}$, not over D, because it requires taking square roots.

Lemma 3.4. — *Let $(D, \dagger)$ be a division $\mathbb{Q}$-algebra with a positive involution of type I or II. Let $\alpha\colon (\mathrm{M}_d(\mathbb{R})^e, t) \to (D_{\mathbb{R}}, \dagger)$ be an isomorphism of $\mathbb{R}$-algebras with involution. Let V be a left D-vector space equipped with a $(D, \dagger)$-skew-Hermitian form $\psi\colon V \times V \to D$. Let $v_1, \dots, v_m$ be a left D-basis for V which is weakly symplectic (when D has type I) or weakly unitary (when D has type II).*

Then there exist $s_1, \dots, s_m \in D_{\mathbb{R}}^{\times}$ such that $s_1^{-1}v_1, \dots, s_m^{-1}v_m$ form a symplectic or α-unitary $D_{\mathbb{R}}$-basis for $V_{\mathbb{R}}$ (according to the type of D) and, for all i,

$$|s_i|_D \leqslant (de)^{1/4} |\psi(v_i, v_j)|_D^{1/2},$$

where j is the unique index such that $\psi(v_i, v_j) \neq 0$.

Proof. — The proof is in two parts, depending on the type of D.

Type I case. — For each $k = 1, \dots, m/2$, $i = 2k - 1$ and $j = 2k$, let

$$t_k = (de)^{-1/2} |\psi(v_i, v_j)|_D \in \mathbb{R}_{>0}.$$

Let $s_i = t_k^{-1/2} \psi(v_i, v_j)$ and $s_j = t_k^{1/2}$. Then

$$\psi(s_i^{-1}v_i, s_j^{-1}v_j) = s_i^{-1}\psi(v_i, v_j)(s_j^{-1})^{\dagger} = 1$$

since $s_j^{\dagger} = s_j$ and $t_k \in \mathbb{R}$ is in the centre of $D_{\mathbb{R}}$.

Furthermore,

$$|s_i|_D = t_k^{-1/2} |\psi(v_i, v_j)|_D = (de)^{1/4} |\psi(v_i, v_j)|_D^{1/2},$$

while

$$|s_j|_D = t_k^{1/2} |1|_D = (de)^{1/2} t_k^{1/2} = (de)^{1/4} |\psi(v_i, v_j)|_D^{1/2}.$$

Type II case. — For each i, $\psi(v_i, v_i) \in D^- \smallsetminus \{0\} \subset F_{\mathbb{R}}^\times \alpha(\omega_0)$. Thus $\psi(v_i, v_i) = t_i\alpha(\omega_0)$ for some $t_i \in F_{\mathbb{R}}^\times$. Write $\alpha^{-1}(t_i) = (t_{i1}, \dots, t_{ie}) \in (\mathbb{R}^\times)^e$. Let $s_i = \alpha(s_{i1}, \dots, s_{ie}) \in D_{\mathbb{R}}^\times$ where $s_{ij} \in \mathbf{GL}_2(\mathbb{R})$ are defined as follows:

$$s_{ij} = \begin{pmatrix} \sqrt{t_{ij}} & 0 \\ 0 & \sqrt{t_{ij}} \end{pmatrix} \quad \text{if } t_{ij} \geqslant 0, \qquad s_{ij} = \begin{pmatrix} \sqrt{-t_{ij}} & 0 \\ 0 & -\sqrt{-t_{ij}} \end{pmatrix} \quad \text{if } t_{ij} < 0.$$

Then

$$\mathrm{Nrd}_{D_{\mathbb{R}}/F_{\mathbb{R}}}(s_i) = \alpha(\det(s_{i1}), \dots, \det(s_{ie})) = \alpha(t_{i1}, \dots, t_{ie}) = t_i.$$

Hence by Lemma 2.11,

$$\psi(s_i^{-1}v_i, s_i^{-1}v_i) = s_i^{-1}\psi(v_i, v_i)(s_i^{-1})^\dagger = \mathrm{Nrd}_{D_{\mathbb{R}}/F_{\mathbb{R}}}(s_i^{-1})\psi(v_i, v_i) = \alpha(\omega_0).$$

Furthermore,

$$\begin{aligned} |s_i|_D^2 &= \sum_{j=1}^{e} \mathrm{Tr}(s_{ij}s_{ij}^t) = \sum_{j=1}^{e} 2|t_{ij}| \\ &\leqslant \sqrt{4e\textstyle\sum_{j=1}^{e}|t_{ij}|^2} = \sqrt{2e\,\mathrm{Trd}_{D_{\mathbb{R}}/\mathbb{R}}(t_i t_i^\dagger)} = (2e)^{1/2}|t_i|_D. \end{aligned}$$

By (3.1), this implies that

$$|s_i|_D^2 \leqslant (2e)^{1/2}|t_i\alpha(\omega_0)|_D = (de)^{1/2}|\psi(v_i, v_i)|_D. \qquad \square$$

Lemma 3.5. — *Let $D_0 = \mathrm{M}_d(k)^e$ where $d = 1$ or 2 and let t denote the involution of D_0 which is transpose on each factor. Let V be a free left D_0-module and let ψ_0 be a non-degenerate (D_0, t)-skew-Hermitian form $V \times V \to D_0$.*

Then there exists a D_0-basis $v_1, \dots, v_m$ for V and a k-basis $a_1, \dots, a_{d^2e}$ for D_0 with the following properties:

(i) *$\{v_1, \dots, v_m\}$ is symplectic with respect to ψ_0 if $d = 1$ and unitary if $d = 2$;*

(ii) *$\{a_1, \dots, a_{d^2e}\}$ is an orthonormal basis for D_0 with respect to $|\cdot|_D$;*

(iii) *$\{a_r v_j : 1 \leqslant r \leqslant d^2e,\ 1 \leqslant j \leqslant m\}$ is a symplectic k-basis for V with respect to $\mathrm{Trd}_{D_0/k}\,\psi_0$.*

Proof. — Write $B_0 = \mathrm{M}_d(k)$. Write F_0 for the centre of D_0, namely k^e. Let $u_1, \dots, u_e$ denote the standard k-basis of $F_0 = k^e$.

Let $V_i = u_i V$. Then $V = \bigoplus_{i=1}^e V_i$ and each V_i is a free left B_0-module. Because V_0 is a free left D_0-module, $\mathrm{rk}_{B_0}(V_1) = \cdots = \mathrm{rk}_{B_0}(V_e)$. Let m denote this rank.

Because $\psi_0 : V \times V \to D_0$ is F_0-bilinear, it takes the form

$$\psi_0((x_1, \dots, x_e), (y_1, \dots, y_e)) = (\psi_1(x_1, y_1), \dots, \psi_e(x_e, y_e)) \quad \text{for all } x_i, y_i \in V_i,$$

where $\psi_i \colon V_i \times V_i \to B_0$ are some non-degenerate (B_0, t)-skew-Hermitian forms.

Below, we shall prove the lemma with (D_0, V, ψ_0) replaced by (B_0, V_i, ψ_i), yielding a B_0-basis $v_{i1}, \dots, v_{im}$ for (V_i, ψ_i) and a k-basis $b_1, \dots, b_{d^2}$ for B_0. Then letting $v_j = (v_{1j}, \dots, v_{ej})$, we obtain a symplectic or unitary D_0-basis for V. Furthermore $\{v_i b_j : 1 \leqslant i \leqslant e, 1 \leqslant j \leqslant d^2\}$ forms a k-basis for D_0 which satisfies (ii) and (iii).

Now we prove the lemma for (B_0, V_i, ψ_i), breaking into two cases depending on d.

Case $d = 1$. — When $d = 1$, $B_0 = k$. Each V_i is a k-vector space of dimension m and ψ_i is a non-degenerate symplectic form $V_i \times V_i \to k$. By the theory of symplectic forms, there exists a symplectic k-basis $\{v_{i1}, \dots, v_{im}\}$ for V_i, proving (i).

Choosing $b_1 = 1$ gives an orthonormal k-basis of B_0 with respect to $|\cdot|_{B_0}$. Since $\operatorname{Trd}_{B_0/k} \psi = \psi$, the bases $v_{i1}, \dots, v_{im}$ and b_1 satisfy (iii).

Case $d = 2$, part (i). — We prove by induction on m that there is a unitary B_0-basis $v_{i1}, \dots, v_{im}$ using the Gram–Schmidt method.

First we claim that there exists $z \in V_i$ such that $\psi_i(z, z) \neq 0$. The image of $\psi_i \colon V_i \times V_i \to B_0$ is a two-sided ideal in B_0, which is a simple algebra, so this image is all of B_0. In particular, we can choose $x, y \in V_i$ such that $\psi_i(x, y)$ is not skew-symmetric, that is, $\psi_i(x, y) + \psi_i(y, x) = \psi_i(x, y) + \psi_i(x, y)^t \neq 0$. Then $\psi_i(x, x)$, $\psi_i(y, y)$ and $\psi_i(x + y, x + y)$ are not all zero. Choosing z to be one of x, y and $x + y$, we obtain $\psi_i(z, z) \neq 0$.

Then $\psi_i(z, z) \in B_0^- = kJ_d$ so $\psi_i(z, z) = sJ_d$ for some $s \in k^\times$. Letting $v_{i1} = \left(\begin{smallmatrix} s^{-1} & 0 \\ 0 & 1 \end{smallmatrix}\right) z$, we obtain that $\psi_i(v_{i1}, v_{i1}) = J_d$.

Let $V_i' = \{v \in V_i : \psi_i(v_{i1}, v) = 0\} = \{v \in V_i : \psi_i(v, v_{i1}) = 0\}$, which is a left B_0-submodule of V_i. For every $b \in B_0 \smallsetminus \{0\}$, we have

$$\psi_i(bv_{i1}, v_{i1}) = b\psi_i(v_{i1}, v_{i1}) = bJ_d \neq 0, \tag{3.2}$$

and so $B_0 v_{i1} \cap V_i' = \{0\}$. For every $v \in V_i$, we have

$$v - \psi_i(v, v_{i1}) J_d^{-1} v_{i1} \in V_i'.$$

Hence $V_i = B_0 v_{i1} \oplus V_i'$ as a direct sum of left B_0-modules.

By (3.2), $bv_{i1} \neq 0$ for all $b \in B_0 \smallsetminus \{0\}$. Hence $\dim_k(B_0 v_{i1}) = 4$ and so $\dim_k(V_0') = 4(m-1)$. Every B_0-module whose k-dimension is a multiple of 4 is a free B_0-module, so $B_0 v_{i1}$ and V_0' are free left B_0-modules. By induction, there is a unitary B_0-basis $v_{i2}, \dots, v_{im}$ for V_i'. Then $v_{i1}, v_{i2}, \dots, v_{im}$ is a unitary B_0-basis for V_i.

Case $d = 2$, part (ii) *and* (iii). — Let

$$b_1 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right), \quad b_2 = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), \quad b_3 = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right), \quad b_4 = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right) \in B_0 = \mathrm{M}_2(k).$$

These form an orthonormal k-basis for B_0 with respect to $|\cdot|_{B_0}$.

Since ψ_i is (B_0, t)-skew-Hermitian,

$$\psi_i(b_r v_{ij}, b_{r'} v_{ij'}) = b_r \psi_i(v_{ij}, v_{ij'}) b_{r'}^\dagger.$$

Thus if $j \neq j'$, we obtain $\psi_i(b_r v_{ij}, b_{r'} v_{ij'}) = 0$. If $j = j'$, then $\psi_i(v_{ij}, v_{ij}) = J_2$, so we can calculate

$$\operatorname{Trd}_{B_0/k} \psi(b_r v_{ij}, b_{r'} v_{ij}) = \operatorname{Trd}_{\mathrm{M}_2(k)/k}(b_r J_2 b_{r'}^t) = \begin{cases} 1 & \text{if } (r, r') = (1, 2) \text{ or } (3, 4), \\ -1 & \text{if } (r, r') = (2, 1) \text{ or } (4, 3), \\ 0 & \text{otherwise.} \end{cases}$$

Thus the bases $v_{i1}, \dots, v_{im}$ and $b_1, \dots, b_4$ satisfy (iii) for (B_0, V_i, ψ_i). □

3.C. Discriminants and skew-Hermitian forms. — The following lemmas are useful for calculating discriminants of skew-Hermitian forms.

Lemma 3.6. — *Let $(D,\dagger)$ be a division $\mathbb{Q}$-algebra with an involution and let R be an order in D. Let $r_1,\dots,r_{d^2e}$ be a $\mathbb{Z}$-basis for R. For $a\in D$, let $T_a\in \mathrm{M}_{d^2e}(\mathbb{Q})$ be the matrix with entries $(T_a)_{ij}=\mathrm{Trd}_{D/\mathbb{Q}}(r_i a r_j^\dagger)$. Then*

$$\det(T_a)=\pm d^{-d^2e}\operatorname{disc}(R)\operatorname{Nm}_{D/\mathbb{Q}}(a).$$

Proof. — Let $M_a\in \mathrm{M}_{d^2e}(\mathbb{Z})$ denote the matrix which represents "multiplication by a on the right" with respect to the basis $r_1,\dots,r_{d^2e}$. Using the facts that $\mathrm{Trd}_{D/\mathbb{Q}}(xy)=\mathrm{Trd}_{D/\mathbb{Q}}(yx)$ for all $x,y\in D$ and that $\mathrm{Trd}_{D/\mathbb{Q}}$ is $\mathbb{Q}$-linear,

$$\begin{aligned}(T_a)_{ij}=\mathrm{Trd}_{D/\mathbb{Q}}(r_i a r_j^\dagger)=\mathrm{Trd}_{D/\mathbb{Q}}(r_j^\dagger r_i a)&=\mathrm{Trd}_{D/\mathbb{Q}}\Big(r_j^\dagger\textstyle\sum_{k=1}^{d^2e}(M_a)_{ki}r_k\Big)\\&=\sum_{k=1}^{d^2e}(M_a)_{ki}\,\mathrm{Trd}_{D/\mathbb{Q}}(r_k r_j^\dagger)=\sum_{k=1}^{d^2e}(M_a)_{ki}(T_1)_{kj}.\end{aligned}$$

Thus $T_a=M_a^tT_1$ so

$$\det(T_a)=\det(M_a)\det(T_1)=\operatorname{Nm}_{D/\mathbb{Q}}(a)\det(T_1).$$

Now T_1 is the Gram matrix of the bilinear form $(x,y)\mapsto d^{-1}\mathrm{Trd}_{D/\mathbb{Q}}(xy^\dagger)$ with respect to $r_1,\dots,r_{d^2e}$. Hence by [DO22, Lem. 5.6], $\det(T_1)=\pm d^{-d^2e}\operatorname{disc}(R)$. □

The following lemma allows us to calculate the discriminant of $\mathrm{Trd}_{D/\mathbb{Q}}\psi$ on a free R-module generated by a weakly symplectic or weakly unitary basis (weakly symplectic or weakly unitary bases with respect to a non-degenerate form automatically satisfy the condition about uniqueness of a permutation σ). We have stated the lemma more generally because we shall require it in one additional case: when $m=2$ and the matrix with entries $\psi(v_i,v_j)$ has the form $\left(\begin{smallmatrix}0&*\\*&*\end{smallmatrix}\right)$.

Lemma 3.7. — *Let $(D,\dagger)$ be a division $\mathbb{Q}$-algebra with a positive involution of type I or II. Let V be a left D-vector space with a non-degenerate $(D,\dagger)$-skew-Hermitian form $\psi\colon V\times V\to D$. Let R be an order in D.*

Let $v_1,\dots,v_m$ be a D-basis for V. Suppose that there is exactly one permutation $\sigma\in S_m$ for which $\psi(v_i,v_{\sigma(i)})\neq 0$ for all $i=1,\dots,m$. Then

$$|\operatorname{disc}(Rv_1+\cdots+Rv_m,\mathrm{Trd}_{D/\mathbb{Q}}\psi)|=d^{-d^2em}|\operatorname{disc}(R)|^m\prod_{i=1}^m|\operatorname{Nm}_{D/\mathbb{Q}}(\psi(v_i,v_{\sigma(i)}))|.$$

Proof. — Choose a $\mathbb{Z}$-basis $r_1,\dots,r_{d^2e}$ for R. Let $A\in \mathrm{M}_n(\mathbb{Q})$ be the Gram matrix of the bilinear form $\mathrm{Trd}_{D/\mathbb{Q}}\psi\colon V\times V\to\mathbb{Q}$ with respect to the $\mathbb{Q}$-basis $r_1v_1,r_2v_1,\dots,r_{d^2e}v_1,\ r_1v_2,\dots,r_{d^2e}v_m$ for V. Then A is made up of square blocks $B_{ij}\in \mathrm{M}_{d^2e}(\mathbb{Q})$ where B_{ij} is the matrix with entries

$$(B_{ij})_{k\ell}=\mathrm{Trd}_{D/\mathbb{Q}}\psi(r_kv_i,r_\ell v_j)=\mathrm{Trd}_{D/\mathbb{Q}}(r_k\psi(v_i,v_j)r_\ell^\dagger).$$

In other words, B_{ij} is equal to the matrix $T_{\psi(v_i,v_j)}$ as defined in Lemma 3.6.

Let $\sigma \in S_m$ be the permutation from the hypothesis of the lemma. By the permutation formula for determinants, the blocks $B_{i\sigma(i)}$ are the only blocks that contribute to $\det(A)$ (although they need not be the only non-zero blocks of A). Indeed, we have

$$\operatorname{disc}(Rv_1 + \cdots + Rv_m, \operatorname{Trd}_{D/\mathbb{Q}} \psi) = \det(A) = \pm \prod_{i=1}^{m} \det(B_{i\sigma(i)}),$$

which, by Lemma 3.6, is equal to

$$\pm \prod_{i=1}^{m} d^{-d^2 e} \operatorname{disc}(R) \operatorname{Nm}_{D/\mathbb{Q}}(\psi(v_i, v_{\sigma(i)})). \qquad \square$$

3.D. *D*-norms. — Let k be a subfield of $\mathbb{R}$. Let $(D, \dagger)$ be a semisimple k-algebra with a positive involution. Let V be a left D-module. We say that a function $|\cdot|\colon V_{\mathbb{R}} \to \mathbb{R}$ is a *D-norm* if it is a norm induced by a positive definite inner product on $V_{\mathbb{R}}$ and it satisfies the inequality

$$|av| \leqslant |a|_D |v| \quad \text{for all } a \in D_{\mathbb{R}},\ x \in V_{\mathbb{R}}.$$

Note that $|\cdot|_D$ is itself a D-norm on $D_{\mathbb{R}}$ thanks to Lemma 2.1.

Let $\psi\colon V \times V \to D$ be a non-degenerate $(D, \dagger)$-skew-Hermitian form. We say that a D-norm $|\cdot|$ is *adapted to* ψ if it satisfies the following two conditions:

(1) $\operatorname{covol}(L_1) = 1$ where $L_1 \subset V$ is the $\mathbb{Z}$-module generated by a symplectic k-basis for V with respect to $\operatorname{Trd}_{D/k} \psi$. (Note that a symplectic basis always exists since $\operatorname{Trd}_{D/k} \psi$ is a symplectic form over a field. Furthermore, this condition is independent of the choice of symplectic k-basis, because the matrix transforming one symplectic basis into another has determinant 1.)

(2) $|\psi(x, y)|_D \leqslant |x||y|$ for all $x, y \in V_{\mathbb{R}}$.

The following two lemmas demonstrate the significance of condition (1) and establish the existence of a D-norm adapted to ψ.

Lemma 3.8. — *Let $(D, \dagger)$ be a division $\mathbb{Q}$-algebra with a positive involution of type I or II. Let V be a left D-vector space with a non-degenerate $(D, \dagger)$-skew-Hermitian form $\psi\colon V \times V \to D$. Let $|\cdot|$ be a D-norm on $V_{\mathbb{R}}$ which satisfies condition* (1) *from the definition of "adapted to ψ." Let L be a $\mathbb{Z}$-lattice in V.*

Then $\operatorname{covol}(L) = |\operatorname{disc}(L)|^{1/2}$, *where we use the volume form associated with $|\cdot|$.*

Proof. — Choose a symplectic $\mathbb{Q}$-basis $e_1, \ldots, e_n$ for V with respect to $\operatorname{Trd}_{D/\mathbb{Q}} \psi$ and a $\mathbb{Z}$-basis $f_1, \ldots, f_n$ for L. Let M be the matrix which maps $e_1, \ldots, e_n$ to $f_1, \ldots, f_n$. The $\mathbb{Z}$-module generated by $e_1, \ldots, e_n$ has covolume 1 by condition (1). Hence $\operatorname{covol}(L) = |\det(M)|$. The matrix with entries $\psi(f_i, f_j)$ is equal to MJ_nM^t. So

$$\operatorname{disc}(L) = \det(MJ_nM^t) = \det(M)^2. \qquad \square$$

Lemma 3.9. — *Let $(D, \dagger)$ be a division $\mathbb{Q}$-algebra with a positive involution of type I or II. Let V be a left D-vector space of dimension m, equipped with a non-degenerate $(D, \dagger)$-skew-Hermitian form $\psi\colon V \times V \to D$. Then there exists a D-norm $|\cdot|$ on $V_{\mathbb{R}}$ which is adapted to ψ.*

Proof. — Identify $D_{\mathbb{R}}$ with $\mathrm{M}_d(\mathbb{R})^e$ where $d = 1$ or 2. By Lemma 3.5(i), there exists a symplectic or unitary $D_{\mathbb{R}}$-basis $v_1, \dots, v_m$ for $V_{\mathbb{R}}$, according to the type of $(D, \dagger)$. Define the following norm on $V_{\mathbb{R}}$:

$$\Big|\sum_{i=1}^m x_i v_i\Big| = \sqrt{\textstyle\sum_{i=1}^m |x_i|_D^2}.$$

This is induced by the inner product $\langle \sum_{i=1}^m x_i v_i, \sum_{j=1}^m y_j v_i \rangle = \mathrm{Trd}_{D_{\mathbb{R}}/\mathbb{R}}\big(\sum_{i=1}^m x_i y_i^\dagger\big)$. It is a D-norm by Lemma 2.1.

Let $a_1, \dots, a_{d^2e}$ be the $\mathbb{R}$-basis for $D_{\mathbb{R}}$ given by Lemma 3.5. Since $a_1, \dots, a_{d^2e}$ is an orthonormal $\mathbb{R}$-basis for $D_{\mathbb{R}}$ with respect to $|\cdot|_D$, $\{a_j v_i\}$ is an orthonormal basis for $V_{\mathbb{R}}$ with respect to $|\cdot|$. Therefore the lattice generated by $\{a_j v_i\}$ has covolume 1. According to Lemma 3.5(iii), $\{a_j v_i\}$ is a symplectic basis for $V_{\mathbb{R}}$ with respect to $\mathrm{Trd}_{D_{\mathbb{R}}/\mathbb{R}} \psi$. Thus the norm $|\cdot|$ satisfies condition (1).

By the triangle inequality for $|\cdot|_D$, we have

$$\Big|\psi\Big(\sum_{i=1}^m x_i v_i, \sum_{j=1}^m y_j v_j\Big)\Big|_D \leqslant \sum_{i=1}^m \sum_{j=1}^m |x_i \psi(v_i, v_j) y_j^\dagger|_D. \tag{3.3}$$

If $\psi(v_i, v_j) \neq 0$, then $\psi(v_i, v_j) = \pm 1$ or ω_0 for all i, j and so by (3.1), $|x_i|_D = |x_i \psi(v_i, v_j)|_D$. Hence

$$|x_i \psi(v_i, v_j) y_j^\dagger|_D \leqslant |x_i \psi(v_i, v_j)|_D |y_j^\dagger|_D = |x_i|_D |y_j|_D. \tag{3.4}$$

Let $\sigma \in S_m$ be the permutation such that $\psi(v_i, v_{\sigma(i)}) \neq 0$ (thus if $(D, \dagger)$ has type I, then $\sigma = (1,2)(3,4)(5,6)\cdots$, while if $(D, \dagger)$ has type II, then $\sigma = \mathrm{id}$). From (3.3) and (3.4), we obtain

$$\Big|\psi\Big(\sum_{i=1}^m x_i v_i, \sum_{j=1}^m y_j v_j\Big)\Big|_D \leqslant \sum_{i=1}^m |x_i|_D |y_{\sigma(i)}|_D.$$

By the Cauchy–Schwarz inequality, we get

$$\Big|\psi\Big(\sum_{i=1}^m x_i v_i, \sum_{j=1}^m y_j v_j\Big)\Big|_D \leqslant \Big(\sum_{i=1}^m |x_i|_D^2\Big)^{1/2} \Big(\sum_{j=1}^m |y_j|_D^2\Big)^{1/2} = \Big|\sum_{i=1}^m x_i v_i\Big|_D \Big|\sum_{j=1}^m y_j v_j\Big|_D.$$

Thus the norm $|\cdot|$ satisfies condition (2). □

4. Proof of Theorem 1.2

In this section we prove our main theorem on weakly unitary or symplectic bases with respect to skew-Hermitian forms. The proof is based on the Gram–Schmidt process, following an inductive structure. For technical reasons we may construct either one or two basis vectors at each step of the induction. Lemma 4.4 constructs the new basis vector(s) for each induction step, and then Proposition 4.5 consists of calculations to keep track of the bounds during this induction.

4.A. Initial vectors of a weakly symplectic or unitary basis. — We would like to begin by choosing v_1 to be the shortest non-zero vector in V (with respect to a suitable D-norm), then inductively choosing a basis for $V^\perp$, the orthogonal complement of Dv_1. However if we do this, $\psi(v_1, v_1)$ might be zero (indeed, if D has type I, then it must be zero) and then $Dv_1 + V^\perp$ is not a direct sum.

We will therefore instead choose either

(1) one short vector $v_1 \in V$ such that $\psi(v_1, v_1) \neq 0$; or

(2) two short vectors $v_1, v_2 \in V$ such that the restriction of ψ to $Dv_1 + Dv_2$ is non-degenerate, and v_1, v_2 form a weakly symplectic or weakly unitary basis for $Dv_1 + Dv_2$.

Let $V^\perp$ denote the orthogonal complement of v_1 (in case (1)) or of $Dv_1 + Dv_2$ (in case (2)). We will bound the discriminant of $\mathrm{Trd}_{D/\mathbb{Q}}\psi$ restricted to $V^\perp$, and then inductively obtain a weakly symplectic or weakly unitary basis for $V^\perp$. Combining this with v_1 and perhaps v_2 gives the basis for V required to prove Theorem 1.2.

The following lemmas choose v_1 and perhaps v_2 satisfying (1) or (2) above.

Lemma 4.1. — *Let $(D, \dagger)$ be a division $\mathbb{Q}$-algebra with an involution. Let V be a left D-vector space, equipped with a non-degenerate $(D, \dagger)$-skew-Hermitian form $\psi\colon V \times V \to D$. Let $w_1, \dots, w_m$ be a D-basis for V. Then there exists a permutation $\sigma \in S_m$ such that $\psi(w_i, w_{\sigma(i)}) \neq 0$ for all $i = 1, \dots, m$.*

Proof. — If D is a field, then the non-degeneracy of ψ implies that the matrix with entries $\psi(w_i, w_j)$ has non-zero determinant. Then the result is immediate by expressing the determinant as an alternating sum over permutations in S_m. When D is non-commutative, we cannot use determinants so we instead use a combinatorial argument (which is also valid in the commutative case).

The argument is based on Hall's theorem on distinct representatives of subsets:

Theorem 4.2 ([Hal35]). — *Let T be a set and let $T_1, \dots, T_m$ be subsets of T. Then there exist pairwise distinct elements $a_1, \dots, a_m$ satisfying $a_i \in T_i$ if and only if, for every $k = 1, \dots, m$ and every choice of k distinct indices $i_1, \dots, i_k$, we have*

$$|T_{i_1} \cup \cdots \cup T_{i_k}| \geqslant k. \tag{4.1}$$

We shall apply Theorem 4.2 with $T = \{1, \dots, m\}$ and

$$T_i = \{j : 1 \leqslant j \leqslant m,\ \psi(w_i, w_j) \neq 0\}.$$

We claim that these sets T_i satisfy the condition (4.1) in Theorem 4.2.

Indeed, suppose that (4.1) is not satisfied for some distinct $i_1, \dots, i_k$. Let W denote the left D-vector space spanned by $w_{i_1}, \dots, w_{i_k}$. Consider the vectors $w \in W$ satisfying

$$\psi(w, w_j) = 0 \text{ for all } j \in T_{i_1} \cup \cdots \cup T_{i_k}. \tag{4.2}$$

Since (4.1) is not satisfied, (4.2) imposes $|T_{i_1} \cup \cdots \cup T_{i_k}| < k = \dim_D(W)$ left D-linear conditions on w. Hence, there exists a non-zero $w \in W$ which satisfies (4.2). By the

definition of the sets T_i and of W, w is also orthogonal to w_j for every $j \notin T_{i_1} \cup \cdots \cup T_{i_k}$. Thus w is orthogonal to all of V. This contradicts the non-degeneracy of ψ.

Hence by Theorem 4.2, there exist pairwise distinct $a_1, \ldots, a_m$ such that $a_i \in T_i$. Since $a_1, \ldots, a_m$ are m distinct elements of $\{1, \ldots, m\}$, the function $\sigma(i) = a_i$ is a permutation of $\{1, \ldots, m\}$. By the definition of T_i, we have $\psi(w_i, w_{\sigma(i)}) \neq 0$ for all i. □

Lemma 4.3. — *Let $(D, \dagger)$ be a division $\mathbb{Q}$-algebra with a positive involution. Let V be a left D-vector space of dimension m, equipped with a non-degenerate $(D, \dagger)$-skew-Hermitian form $\psi\colon V \times V \to D$. Let $|\cdot|$ be a D-norm on $V_{\mathbb{R}}$. Let $w_1, \ldots, w_m$ be a D-basis for V.*

Then there exist $i, j \in \{1, \ldots, m\}$ satisfying the following conditions:

(i) $|w_i||w_j| \leqslant \big(|w_1||w_2| \cdots |w_m|\big)^{2/m}$;
(ii) $\psi(w_i, w_j) \neq 0$;
(iii) *if $i \neq j$, then $\psi(w_i, w_i) = 0$.*

Proof. — Let σ be a permutation as in Lemma 4.1.

Choose $k \in \{1, \ldots, m\}$ so that $|w_k||w_{\sigma(k)}|$ is minimal. Then

$$|w_k||w_{\sigma(k)}| \leqslant \Big(\prod_{i=1}^{m} |w_i||w_{\sigma(i)}|\Big)^{1/m} = \Big(\prod_{i=1}^{m} |w_i| \cdot \prod_{j=1}^{m} |w_j|\Big)^{1/m} = \Big(\prod_{i=1}^{m} |w_i|\Big)^{2/m}.$$

By the choice of σ, we have $\psi(w_k, w_{\sigma(k)}) \neq 0$.

– If $\sigma(k) = k$, then $i = j = k$ satisfies the conditions of the lemma.
– Otherwise choose $i \in \{k, \sigma(k)\}$ so that $|w_i|$ is minimal.
– If $\psi(w_i, w_i) \neq 0$, then choosing $j = i$ satisfies the required conditions.
– If $\psi(w_i, w_i) = 0$, then choose j to be the element of $\{k, \sigma(k)\}$ which is different from i. This i and j satisfy the required conditions. □

In the remainder of this section, whenever we refer to a discriminant other than $\mathrm{disc}(R)$, we mean the discriminant of $\mathrm{Trd}_{D/\mathbb{Q}}\,\psi$ restricted to the specified $\mathbb{Z}$-module.

Lemma 4.4. — *Let $(D, \dagger)$ be a division $\mathbb{Q}$-algebra with a positive involution of type I or II. Let V be a left D-vector space with a non-degenerate $(D, \dagger)$-skew-Hermitian form $\psi\colon V \times V \to D$. Let L be a $\mathbb{Z}$-lattice in V such that $\mathrm{Trd}_{D/\mathbb{Q}}\,\psi(L \times L) \subset \mathbb{Z}$. Let R be an order which is contained in $\mathrm{Stab}_D(L)$ and let $\eta \in \mathbb{Z}_{>0}$ be a positive integer such that $\eta R^\dagger \subset R$.*

Then there exists an R-submodule $M \subset L$ with the following properties:

(i) $r := \dim_D(D \otimes_R M) = 1$ *or* 2;
(ii) *the restriction of ψ to M is non-degenerate;*
(iii) $|\mathrm{disc}(M)| \leqslant (\gamma_{d^2em}^2/d^3e)^{d^2er/2} |\mathrm{disc}(R)|^r |\mathrm{disc}(L)|^{r/m}$;
(iv) *one of the following occurs:*

(a) *D has type I, $r = 2$ and $M = Rv_1 + Rv_2$ for some v_1, v_2 such that*

$$|\psi(v_1, v_2)|_D \leqslant \gamma_{em} |\mathrm{disc}(L)|^{1/em};$$

(b) *D has type II, $r = 1$ and $M = Rv_1$ for some v_1 such that*

$$|\psi(v_1, v_1)|_D \leqslant \gamma_{4em} |\mathrm{disc}(L)|^{1/4em};$$

(c) *D has type II, $r = 2$ and there exist D-linearly independent vectors $v_1, v_2 \in M$ such that $\psi(v_1, v_2) = 0$,*

$$|\psi(v_1, v_1)|_D, |\psi(v_2, v_2)|_D \leqslant 2^{-5/2} \gamma_e^{1/2} \gamma_{4em}^2 \eta^7 |\mathrm{disc}(R)|^{2/e} |\mathrm{disc}(L)|^{1/2em},$$

and

$$[M : Rv_1 + Rv_2] \leqslant (\gamma_e/8e)^{2e} (\gamma_{4em}^2/8e)^{2e} \eta^{28e} |\mathrm{disc}(R)|^8 |\mathrm{disc}(L)|^{1/m}.$$

Proof. — By Lemma 3.9, there is a D-norm $|\cdot|$ on $V_{\mathbb{R}}$ adapted to ψ. By Proposition 2.5, there exists a D-basis $w_1, \dots, w_m$ for V satisfying $w_1, \dots, w_m \in L$ and

$$|w_1| \cdots |w_m| \leqslant \gamma_{d^2em}^{m/2} \operatorname{covol}(L)^{1/d^2e} \leqslant \gamma_{d^2em}^{m/2} |\mathrm{disc}(L)|^{1/2d^2e},$$

where the second inequality comes from Lemma 3.8.

Choose i, j as in Lemma 4.3. Since $|\cdot|$ is adapted to ψ, we have

$$|\psi(w_i, w_j)|_D \leqslant |w_i||w_j| \leqslant \gamma_{d^2em} |\mathrm{disc}(L)|^{1/d^2em}. \tag{4.3}$$

Proof of (i)–(iii). — Let $M = Rw_i + Rw_j$, so that $r = 1$ if $i = j$ and $r = 2$ if $i \neq j$.

If $i = j$, then by Lemma 4.3, $\psi(w_i, w_i) \neq 0$, so the restriction of ψ to M is non-degenerate.

If $i \neq j$, then by Lemma 4.3, $\psi(w_i, w_i) = 0$ and $\psi(w_i, w_j) \neq 0$. Consequently, for any vector $x \in M$, if $x \in Dw_i \smallsetminus \{0\}$ then $\psi(x, w_j) \neq 0$ while if $x \notin Dw_i$ then $\psi(x, w_i) \neq 0$. Thus the restriction of ψ to M is non-degenerate.

By Lemma 3.7, Lemma 2.2 and (4.3), we obtain that in both cases $i = j$ or $i \neq j$,

$$\begin{aligned}
|\mathrm{disc}(M)| &= d^{-d^2er} |\mathrm{disc}(R)|^r |\mathrm{Nm}_{D/\mathbb{Q}}(\psi(w_i, w_j))|^r \\
&= d^{-d^2er} |\mathrm{disc}(R)|^r |\mathrm{Nrd}_{D/\mathbb{Q}}(\psi(w_i, w_j))|^{dr} \\
&\leqslant d^{-d^2er} |\mathrm{disc}(R)|^r (de)^{-d^2er/2} |\psi(w_i, w_j)|_D^{d^2er} \\
&\leqslant (d^3e)^{-d^2er/2} |\mathrm{disc}(R)|^r \cdot \gamma_{d^2em}^{d^2er} |\mathrm{disc}(L)|^{r/m}.
\end{aligned}$$

For the proof of (iv), we split into cases depending on the type of D and on whether $i = j$ or $i \neq j$.

Case (a). — If D has type I, then D is a field and ψ is a symplectic form. Hence $\psi(v, v) = 0$ for all $v \in V$, so we must have $i \neq j$.

Let $v_1 = w_i$ and $v_2 = w_j$. The bound in (iv)(a) is (4.3).

Case (b). — If D has type II and $i = j$, then let $v_1 = w_i$. Then (iv)(b) holds thanks to (4.3).

Case (c). — If D has type II and $i \neq j$, then choose $\omega \in D^-$ as in Lemma 2.12. Let

$$w_j' = 2\psi(w_i, w_j)\omega w_j - \omega \psi(w_j, w_j) w_i.$$

Since $\mathrm{Trd}_{D/\mathbb{Q}}\,\psi(L\times L)\subset\mathbb{Z}$, $\psi(L\times L)\subset R^*$. Hence $\psi(w_i,w_j)\omega$ and $\omega\psi(w_j,w_j)\in R$, so $w_j'\in Rw_i+Rw_j=M$. Furthermore w_j' and w_i are D-linearly independent because $\psi(w_i,w_j)\omega\neq 0$.

By Lemma 2.11(i), $\omega\psi(w_j,w_j),\psi(w_j,w_j)\omega\in F$. Using this, along with the facts that $\psi(w_i,w_i)=0$ and $(\omega\psi(w_i,w_j))^\dagger=\psi(w_j,w_i)\omega$, we can calculate

$$\begin{aligned}\psi(w_j',w_j') &= 2\psi(w_i,w_j)\omega\,\psi(w_j,w_j)\,(2\psi(w_i,w_j)\omega)^\dagger\\ &\quad -2\psi(w_i,w_j)\omega\,\psi(w_j,w_i)\,(\omega\psi(w_j,w_j))^\dagger\\ &\quad -\omega\psi(w_j,w_j)\,\psi(w_i,w_j)\,(2\psi(w_i,w_j)\omega)^\dagger+0\\ &= (4-2-2)\psi(w_i,w_j)\omega\psi(w_j,w_j)\omega\psi(w_j,w_i)\\ &= 0.\end{aligned}$$

Using Lemma 2.11(ii) and the fact that $\psi(w_i,w_i)=0$, we can calculate

$$\psi(w_j',w_i)=2\psi(w_i,w_j)\omega\,\psi(w_j,w_i)-0=-2\,\mathrm{Nrd}_{D/F}(\psi(w_i,w_j))\omega.$$

Thus $\psi(w_j',w_i)\in F\omega=D^-$, so $\psi(w_i,w_j')=-\psi(w_j',w_i)^\dagger=\psi(w_j',w_i)$.

Now let

$$v_1=w_i-w_j',\qquad v_2=w_i+w_j'.$$

Clearly $v_1,v_2\in Rw_i+Rw_j'\subset M$. Since $w_i=\frac{1}{2}(v_1+v_2)$ and $w_j'=\frac{1}{2}(v_2-v_1)$, the vectors v_1 and v_2 are D-linearly independent.

Since $\psi(w_j',w_i)=\psi(w_i,w_j')$ we can calculate

$$\begin{aligned}&\psi(v_1,v_2)=\psi(w_i,w_i)+\psi(w_i,w_j')-\psi(w_j',w_i)-\psi(w_j',w_j')=0,\\ &\psi(v_1,v_1)=\psi(w_i,w_i)-\psi(w_i,w_j')-\psi(w_j',w_i)+\psi(w_j',w_j')=-2\psi(w_j',w_i),\\ &\psi(v_2,v_2)=\psi(w_i,w_i)+\psi(w_i,w_j')+\psi(w_j',w_i)+\psi(w_j',w_j')=2\psi(w_j',w_i).\end{aligned}$$

Consequently, using Lemmas 2.1, 2.2 and 2.12 and (4.3),

$$\begin{aligned}|\psi(v_1,v_1)|_D=|\psi(v_2,v_2)|_D=2|\psi(w_j',w_i)|_D &\leqslant 4|\mathrm{Nrd}_{D/F}(\psi(w_i,w_j))|_D|\omega|_D\\ &\leqslant 4\cdot 2^{-1/2}|\psi(w_i,w_j)|_D^2\cdot 2^{-4}\eta^7\sqrt{\gamma_e}|\mathrm{disc}(R)|^{2/e}\\ &= 2^{-5/2}\sqrt{\gamma_e}\eta^7|\mathrm{disc}(R)|^{2/e}\cdot\gamma_{4em}^2|\mathrm{disc}(L)|^{2/4em}.\end{aligned}$$

This proves the first inequality in (iv)(c).

Using Lemma 3.7, we have

$$\begin{aligned}[M:Rv_1+Rv_2] &= \frac{|\mathrm{disc}(Rv_1+Rv_2)|^{1/2}}{|\mathrm{disc}(M)|^{1/2}}\\ &= \frac{|\mathrm{Nm}_{D/\mathbb{Q}}(\psi(v_1,v_1))|^{1/2}|\mathrm{Nm}_{D/\mathbb{Q}}(\psi(v_2,v_2))|^{1/2}}{|\mathrm{Nm}_{D/\mathbb{Q}}(\psi(w_i,w_j))|^{1/2}|\mathrm{Nm}_{D/\mathbb{Q}}(\psi(w_j,w_i))|^{1/2}}\\ &= \frac{|\mathrm{Nrd}_{D/\mathbb{Q}}(\psi(v_1,v_1))||\mathrm{Nrd}_{D/\mathbb{Q}}(\psi(v_2,v_2))|}{|\mathrm{Nrd}_{D/\mathbb{Q}}(\psi(w_i,w_j))|^2}.\end{aligned}$$

Now by Lemma 2.2 and the fact that if $a \in F$, then $\mathrm{Nrd}_{D/\mathbb{Q}}(a) = \mathrm{Nm}_{F/\mathbb{Q}}(a)^2$,

$$\begin{aligned}|\mathrm{Nrd}_{D/\mathbb{Q}}(\psi(v_1,v_1))| &= |\mathrm{Nrd}_{D/\mathbb{Q}}(\psi(v_2,v_2))| \\ &= |\mathrm{Nrd}_{D/\mathbb{Q}}(4\,\mathrm{Nrd}_{D/F}(\psi(w_i,w_j))\omega)| \\ &= 4^{2e}|\mathrm{Nm}_{F/\mathbb{Q}}(\mathrm{Nrd}_{D/F}(\psi(w_i,w_j))|^2|\mathrm{Nrd}_{D/\mathbb{Q}}(\omega)| \\ &= 4^{2e}|\mathrm{Nrd}_{D/\mathbb{Q}}(\omega)||\mathrm{Nrd}_{D/\mathbb{Q}}(\psi(w_i,w_j))|^2.\end{aligned}$$

Therefore by Lemmas 2.2 and 2.12 and (4.3),

$$\begin{aligned}[M : Rv_1 + Rv_2] &= \frac{4^{4e}|\mathrm{Nrd}_{D/\mathbb{Q}}(\omega)|^2|\mathrm{Nrd}_{D/\mathbb{Q}}(\psi(w_i,w_j))|^4}{|\mathrm{Nrd}_{D/\mathbb{Q}}(\psi(w_i,w_j))|^2} \\ &= 4^{4e}|\mathrm{Nrd}_{D/\mathbb{Q}}(\omega)|^2|\mathrm{Nrd}_{D/\mathbb{Q}}(\psi(w_i,w_j))|^2 \\ &\leqslant 4^{4e}\cdot(2e)^{-2e}|\omega|_D^{4e}\cdot(2e)^{-2e}|\psi(w_i,w_j)|_D^{4e} \\ &\leqslant 2^{4e}e^{-4e}\cdot 2^{-16e}\eta^{28e}\gamma_e^{2e}|\mathrm{disc}(R)|^8\cdot\gamma_{4em}^{4e}|\mathrm{disc}(L)|^{4e/4em}. \qquad \square\end{aligned}$$

4.B. Inductive construction of weakly symplectic or unitary basis. — The following theorem is a slight generalisation of Theorem 1.2, together with explicit values for the constants. Compared to Theorem 1.2, we only require $R \subset \mathrm{Stab}_D(L)$ (allowing $R \subsetneqq \mathrm{Stab}_D(L)$ is needed for the induction) and we add an additional parameter η. When $R = \mathrm{Stab}_D(L)$, the parameter η is controlled by Lemma 4.6.

Proposition 4.5. — *Let $(D,\dagger)$ be a division $\mathbb{Q}$-algebra with a positive involution of type I or II. Let V be a left D-vector space with a non-degenerate $(D,\dagger)$-skew-Hermitian form $\psi\colon V\times V\to D$. Let L be a $\mathbb{Z}$-lattice in V such that $\mathrm{Trd}_{D/\mathbb{Q}}\,\psi(L\times L)\subset\mathbb{Z}$. Let R be an order which is contained in $\mathrm{Stab}_D(L)$ and let $\eta\in\mathbb{Z}_{>0}$ be a positive integer such that $\eta R^\dagger\subset R$.*

Then there exists a D-basis $v_1,\dots,v_m$ for V such that:

(i) *$v_1,\dots,v_m\in L$;*

(ii) *the basis is weakly symplectic (when D has type I) or weakly unitary (when D has type II) with respect to ψ;*

(iii) *the index of $Rv_1+\cdots+Rv_m$ in L is bounded as follows:*

$$[L : Rv_1+\cdots+Rv_m] \leqslant C_{10}(d,e,m)\eta^{C_{11}(d,e,m)}|\mathrm{disc}(R)|^{C_{12}(d,e,m)}|\mathrm{disc}(L)|^{C_{13}(d,e,m)};$$

(iv) *for all $i,j\in\{1,\dots,m\}$ such that $\psi(v_i,v_j)\neq 0$,*

$$|\psi(v_i,v_j)|_D \leqslant C_{14}(d,e,m)\eta^{C_{15}(d,e,m)}|\mathrm{disc}(R)|^{C_{16}(d,e,m)}|\mathrm{disc}(L)|^{C_{17}(d,e,m)}.$$

The inequalities (iii) *and* (iv) *hold with the values of the constants given by Table* 1.

Proof. — The proof is by induction on $m = \dim_D(V)$. Let M be an R-submodule of L as in Lemma 4.4. Let $r = \dim_D(D\otimes_R M) = 1$ or 2. Choose v_1 and perhaps v_2 as in Lemma 4.4(iv).

For part (iii), the base case of the induction will be when $m = r$, and this is dealt with in the three cases below. For part (iv), the base case is when $m = 0$, in which case (iv) is vacuously true.

Table 1.

	$d=1$	$d=2$
$C_{10}(d,e,m)$	$(em^2)^{em(m+2)/16}$	$(2em^2)^{em(m+2)/2}$
$C_{11}(d,e,m)$	0	$14em$
$C_{12}(d,e,m)$	$m(m+2)/8$	$m(m+16)/4$
$C_{13}(d,e,m)$	$(m-2)/4$	$(m-1)/2$
$C_{14}(d,e,m)$	$(em^2)^{(m(m+2)+24)/32}$	$(2em^2)^{(m(m+1)+14)/8}$
$C_{15}(d,e,m)$	0	7
$C_{16}(d,e,m)$	$\big(m(m+2)-8\big)/16e$	$\big(m(m+1)+26\big)/16e$
$C_{17}(d,e,m)$	$(m+2)/8e$	$(m+1)/8e$

Let $M^\perp$ be the orthogonal complement of M in L with respect to ψ. By Lemma 3.2, $M^\perp$ is also the orthogonal complement of M in L with respect to $\mathrm{Trd}_{D/\mathbb{Q}}\psi$. By Lemma 2.6 and Lemma 4.4(iii),

$$\begin{aligned}|\mathrm{disc}(M^\perp)| &\leqslant |\mathrm{disc}(L)|\cdot|\mathrm{disc}(M)|\\ &\leqslant (\gamma_{d^2em}^2/d^3e)^{d^2er/2}|\mathrm{disc}(R)|^r|\mathrm{disc}(L)|^{(m+r)/m}.\end{aligned}\tag{4.4}$$

Now ψ restricted to $M^\perp$ is non-degenerate, $\dim_D(D\otimes_R M^\perp)=m-r<m$ and $R\subset \mathrm{Stab}_D(M^\perp)$ so we can apply the lemma inductively to $M^\perp$. We obtain a D-basis $v_{r+1},\dots,v_m$ for $D\otimes_R M^\perp$ whose elements lie in $M^\perp\subset L$.

Now $v_1,\dots,v_r\in M$ are orthogonal to $v_{r+1},\dots,v_m$ and $v_1,\dots,v_r$ form a weakly symplectic or weakly unitary D-basis for $D\otimes_R M$. Hence by induction $v_1,\dots,v_m$ form a weakly symplectic or weakly unitary D-basis for V. Thus (i) and (ii) are satisfied.

By induction,

$$\begin{aligned}&[M^\perp : Rv_{r+1}+\cdots+Rv_m]\\ &\qquad\leqslant C_{10}(d,e,m-r)\eta^{C_{11}(d,e,m-r)}|\mathrm{disc}(R)|^{C_{12}(d,e,m-r)}\\ &\qquad\qquad\cdot|\mathrm{disc}(M^\perp)|^{C_{13}(d,e,m-r)}\\ &\qquad\leqslant C_{10}(d,e,m-r)\eta^{C_{11}(d,e,m-r)}|\mathrm{disc}(R)|^{C_{12}(d,e,m-r)}\\ &\qquad\qquad\cdot(\gamma_{d^2em}^2/d^3e)^{d^2er/2\cdot C_{13}(d,e,m-r)}|\mathrm{disc}(R)|^{rC_{13}(d,e,m-r)}\\ &\qquad\qquad\cdot|\mathrm{disc}(L)|^{(m+r)/m\cdot C_{13}(d,e,m-r)}.\end{aligned}\tag{4.5}$$

We now split into cases depending on the type of D and on whether $r=1$ or 2, as in Lemma 4.4(iv). The proofs in the three cases are very similar, with just the details of the calculations varying. For each case, the proofs of (iii) and (iv) are independent of each other.

Case (a), part (iii). — This is the case when D has type I and $r=2$.

When $m=r=2$, from Lemma 4.4(iii) and (iv)(a), we have

$$[L:Rv_1+Rv_2]=[L:M]=\frac{|\mathrm{disc}(M)|^{1/2}}{|\mathrm{disc}(L)|^{1/2}}\leqslant(\gamma_{2e}^2/e)^{e/2}|\mathrm{disc}(R)|.$$

This establishes (iii) when $m = 2$ because

$$(\gamma_{2e}^2/e)^{e/2} \leqslant (4e)^{e/2} = C_{10}(1,e,2),$$
$$C_{11}(1,e,2) = 0, \quad C_{12}(1,e,2) = 1, \quad C_{13}(1,e,2) = 0.$$

When $m \geqslant 3$, we have, using the fact that $M = Rv_1 + Rv_2$, along with Lemma 2.6, Lemma 4.4(iii) and (4.5),

$$\begin{aligned}[L : Rv_1 + \cdots + Rv_m] &= [L : M + M^\perp][M^\perp : Rv_3 + \cdots + Rv_m] \\ &\leqslant |\mathrm{disc}(M)|[M^\perp : Rv_3 + \cdots + Rv_m] \\ &\leqslant (\gamma_{em}^2/e)^e |\mathrm{disc}(R)|^2 |\mathrm{disc}(L)|^{2/m} \cdot C_{10}(1,e,m-2)\,(\gamma_{em}^2/e)^{eC_{13}(1,e,m-2)} \\ &\qquad \cdot |\mathrm{disc}(R)|^{C_{12}(1,e,m-2)+2C_{13}(1,e,m-2)} |\mathrm{disc}(L)|^{(m+2)/m \cdot C_{13}(1,e,m-2)}.\end{aligned}$$

Now we can calculate: for the multiplicative constant:

$$\begin{aligned}C_{10}(1,e,m-2)\,(\gamma_{em}^2/e)^{e(1+C_{13}(1,e,m-2))} &= (e(m-2)^2)^{e(m-2)m/16} (\gamma_{em}^2/e)^{em/4} \\ &\leqslant (em^2)^{e(m-2)m/16} \cdot (em^2)^{m/4} \\ &= (em^2)^{e(m^2-2m+4m)/16} \\ &= C_{10}(1,e,m),\end{aligned}$$

for the exponent of $|\mathrm{disc}(R)|$:

$$\begin{aligned}2 + C_{12}(1,e,m-2) + 2C_{13}(1,e,m-2) &= 2 + \frac{(m-2)m}{8} + 2 \cdot \frac{m-4}{4} \\ &= C_{12}(1,e,m),\end{aligned}$$

for the exponent of $|\mathrm{disc}(L)|$:

$$\begin{aligned}\frac{2}{m} + \frac{(m+2)}{m} \cdot C_{13}(1,e,m-2) &= \frac{2}{m} + \frac{(m+2)(m-4)}{4m} \\ &= \frac{8 + (m^2 - 2m - 8)}{4m} = \frac{(m-2)m}{4m} = C_{13}(1,e,m).\end{aligned}$$

Case (a), part (iv). — For $i = 1$, $j = 2$, Lemma 4.4(iv)(a) gives

$$|\psi(v_1, v_2)|_D \leqslant \gamma_{em} |\mathrm{disc}(L)|^{1/em}. \tag{4.6}$$

This establishes (iv) when $i = 1$, $j = 2$ because, using Lemma 2.3 and the fact that $m \geqslant 2$ so $1 \leqslant (m(m+2)+24)/32$ and $1 \leqslant (m(m+2)+8)/16$,

$$\begin{aligned}\gamma_{em} &\leqslant em \leqslant C_{14}(1,e,m), \\ 0 &\leqslant \frac{m(m+2)-8}{16e} = C_{16}(1,e,m), \\ \frac{1}{em} &\leqslant \frac{2 \cdot 4}{8em} \leqslant \frac{m(m+2)}{8em} = C_{17}(1,e,m).\end{aligned}$$

For $i, j \geqslant 3$, induction gives

$$\begin{aligned}|\psi(v_i, v_j)|_D &\leqslant C_{14}(1, e, m-2)|\mathrm{disc}(R)|^{C_{16}(1,e,m-2)} \mathrm{disc}(M^\perp)^{C_{17}(1,e,m)} \\ &\leqslant C_{14}(1, e, m-2)|\mathrm{disc}(R)|^{C_{16}(1,e,m-2)} \\ &\qquad \cdot \big((\gamma_{em}^2/e)^e |\mathrm{disc}(R)|^2 |\mathrm{disc}(L)|^{(m+2)/m}\big)^{C_{17}(1,e,m-2)}.\end{aligned}$$

Now we can calculate: for the multiplicative constant (using Lemma 2.3):

$$\begin{aligned}C_{14}(1, e, m-2)\,(\gamma_{em}^2/e)^{eC_{17}(1,e,m-2)} &= (e(m-2)^2)^{((m-2)m+24)/32}\,(\gamma_{em}^2/e)^{m/8} \\ &\leqslant (em^2)^{((m-2)m+24)/32} \cdot (em^2)^{m/8} \\ &= (em^2)^{(m^2-2m+24+4m)/32} \\ &= C_{14}(1, e, m),\end{aligned}$$

for the exponent of $|\mathrm{disc}(R)|$:

$$C_{16}(1, e, m-2) + 2C_{17}(1, e, m-2) = \frac{(m-2)m-8}{16e} + 2 \cdot \frac{m}{8e} = C_{16}(1, e, m),$$

for the exponent of $|\mathrm{disc}(L)|$:

$$\frac{m+2}{m} C_{17}(1, e, m-2) = \frac{(m+2)}{m} \cdot \frac{m}{8e} = C_{17}(1, e, m).$$

Case (b), part (iii). — In this case, D has type II and $r = 1$.

When $m = r = 1$, from Lemma 4.4(iii) and (iv)(b), we have

$$[L : Rv_1] = [L : M] = \frac{|\mathrm{disc}(M)|^{1/2}}{|\mathrm{disc}(L)|^{1/2}} \leqslant (\gamma_{4e}^2/8e)^e |\mathrm{disc}(R)|^{1/2}.$$

This establishes (iii) when $m = 1$ because

$$(\gamma_{4e}^2/8e)^e \leqslant (2e)^e \leqslant (2e)^{3e/2} = C_{10}(2, e, 1),$$
$$C_{11}(2, e, 1) = 14e > 0, \quad C_{12}(2, e, 1) = 17/4 > 1/2, \quad C_{13}(2, e, 1) = 0.$$

When $m \geqslant 2$, we have (using $M = Rv_1$, Lemma 2.6, Lemma 4.4(iii) and (4.5))

$$\begin{aligned}[L : Rv_1 + \cdots + Rv_m] &= [L : M + M^\perp][M^\perp : Rv_2 + \cdots + Rv_m] \\ &\leqslant |\mathrm{disc}(M)|[M^\perp : Rv_2 + \cdots + Rv_m] \\ &\leqslant (\gamma_{4em}^2/8e)^{2e} |\mathrm{disc}(R)||\mathrm{disc}(L)|^{1/m} \\ &\qquad \cdot C_{10}(2, e, m-1)\eta^{C_{11}(2,e,m-1)} (\gamma_{4em}^2/8e)^{2eC_{13}(2,e,m-1)} \\ &\qquad \cdot |\mathrm{disc}(R)|^{C_{12}(2,e,m-1)+C_{13}(2,e,m-1)} |\mathrm{disc}(L)|^{(m+1)/m \cdot C_{13}(2,e,m-1)}.\end{aligned}$$

Now we can calculate: for the multiplicative constant:

$$\begin{aligned}C_{10}(2, e, m-1)(\gamma_{4em}^2/8e)^{2e(1+C_{13}(2,e,m-1))} &= (2e(m-1)^2)^{e(m-1)(m+1)/2} (\gamma_{4em}^2/8e)^{em} \\ &\leqslant (2em^2)^{e(m-1)(m+1)/2} \cdot (2em^2)^{em} \\ &= (2em^2)^{e(m^2-1+2m)/2} \\ &\leqslant C_{10}(2, e, m),\end{aligned}$$

for the exponent of η:

$$C_{11}(2,e,m-1) = 14e(m-1) < C_{11}(2,e,m),$$

for the exponent of $|\mathrm{disc}(R)|$:

$$\begin{aligned} 1 + C_{12}(2,e,m-1) + C_{13}(2,e,m-1) &= 1 + \frac{(m-1)(m+15)}{4} + \frac{m-2}{2} \\ &= \frac{m^2+16m-15}{4} < C_{12}(2,e,m), \end{aligned}$$

for the exponent of $|\mathrm{disc}(L)|$:

$$\begin{aligned} \frac{1}{m} + \frac{(m+1)}{m} \cdot C_{13}(2,e,m-1) &= \frac{1}{m} + \frac{(m+1)(m-2)}{2m} \\ &= \frac{2+(m^2-m-2)}{2m} = \frac{(m-1)m}{2m} = C_{13}(2,e,m). \end{aligned}$$

Case (b), part (iv). — For $i = j = 1$, Lemma 4.4(iv)(b) gives

$$|\psi(v_1,v_1)|_D \leqslant \gamma_{4em} |\mathrm{disc}(L)|^{1/4em}. \tag{4.7}$$

This establishes (iv) for $i = j = 1$ because, using Lemma 2.3 and the fact that $m \geqslant 1$, so $(m(m+1)+14)/8 \geqslant 2$,

$$\begin{aligned} \gamma_{4em} &\leqslant 4em \leqslant (2em^2)^2 \leqslant C_{14}(2,e,m), \\ 0 &< 7 = C_{15}(2,e,m), \\ 0 &< \frac{1\cdot 2 + 26}{16e} \leqslant \frac{m(m+1)+26}{16e} = C_{16}(2,e,m), \\ \frac{1}{4em} &\leqslant \frac{2}{8e} \leqslant \frac{m+1}{8e} = C_{17}(2,e,m). \end{aligned}$$

For $i = j \geqslant 2$, induction gives

$$\begin{aligned} |\psi(v_j,v_j)|_D &\leqslant C_{14}(2,e,m-1)\eta^{C_{15}(2,e,m-1)} |\mathrm{disc}(R)|^{C_{16}(2,e,m-1)} \\ &\quad \cdot |\mathrm{disc}(M^\perp)|^{C_{17}(2,e,m-1)} \\ &\leqslant C_{14}(2,e,m-1)\eta^{C_{15}(2,e,m-1)} |\mathrm{disc}(R)|^{C_{16}(2,e,m-1)} \\ &\quad \cdot \big((\gamma_{4em}^2/8e)^{2e} |\mathrm{disc}(R)||\mathrm{disc}(L)|^{(m+1)/m}\big)^{C_{17}(2,e,m-1)}. \end{aligned}$$

Now we can calculate: for the multiplicative constant:

$$\begin{aligned} C_{14}(2,e,m-1)(\gamma_{4em}^2/8e)^{2eC_{17}(2,e,m-1)} &= (2e(m-1)^2)^{((m-1)m+14)/8} \cdot (\gamma_{4em}^2/8e)^{m/4} \\ &\leqslant (2em^2)^{(m^2-m+14)/8} \cdot (2em^2)^{2m/8} \\ &= (2em^2)^{(m^2+m+14)/8} \\ &= C_{14}(2,e,m), \end{aligned}$$

for the exponent of η:

$$C_{15}(2,e,m-1) = 7 = C_{15}(2,e,m),$$

for the exponent of $|\mathrm{disc}(R)|$:

$$C_{16}(2,e,m-1)+C_{17}(2,e,m-1)=\frac{(m-1)m+26}{16e}+\frac{m}{8e}=C_{16}(2,e,m),$$

for the exponent of $|\mathrm{disc}(L)|$:

$$\frac{m+1}{m}C_{17}(2,e,m-1)=\frac{(m+1)}{m}\cdot\frac{m}{8e}=\frac{m+1}{8e}=C_{17}(2,e,m).$$

Case (c), part (iii). — This is the case where D has type II and $r=2$.

When $m=r=2$, from Lemma 4.4(iii) and (iv)(c), we have

$$\begin{aligned}
[L:Rv_1+Rv_2]&=[L:M][M:Rv_1+Rv_2]\\
&=\frac{|\mathrm{disc}(M)|^{1/2}}{|\mathrm{disc}(L)|^{1/2}}[M:Rv_1+Rv_2]\\
&\leqslant\frac{(\gamma_{8e}^2/8e)^{2e}|\mathrm{disc}(R)||\mathrm{disc}(L)|^{1/2}}{|\mathrm{disc}(L)|^{1/2}}(\gamma_e/8e)^{2e}(\gamma_{8e}^2/8e)^{2e}\eta^{28e}|\mathrm{disc}(R)|^8|\mathrm{disc}(L)|^{1/2}\\
&=(\gamma_e/8e)^{2e}(\gamma_{8e}^2/8e)^{4e}\eta^{28e}|\mathrm{disc}(R)|^9|\mathrm{disc}(L)|^{1/2}.
\end{aligned}$$

This establishes (iii) when $m=2$ because

$$(\gamma_e/8e)^{2e}(\gamma_{8e}^2/8e)^{4e}\leqslant 1\cdot(8e)^{4e}=C_{10}(2,e,2),$$
$$C_{11}(2,e,2)=28e,\quad C_{12}(2,e,2)=9,\quad C_{13}(2,e,2)=1/2.$$

When $m\geqslant 3$, we have (using Lemma 2.6, Lemma 4.4(iv)(c) and (4.5))

$$\begin{aligned}
&[L:Rv_1+\cdots+Rv_m]\\
&\quad=[L:M+M^\perp][M:Rv_1+Rv_2][M^\perp:Rv_3+\cdots+Rv_m]\\
&\quad\leqslant|\mathrm{disc}(M)|[M:Rv_1+Rv_2][M^\perp:Rv_3+\cdots+Rv_m]\\
&\quad\leqslant(\gamma_{4em}^2/8e)^{4e}|\mathrm{disc}(R)|^2|\mathrm{disc}(L)|^{2/m}\\
&\qquad\cdot(\gamma_e/8e)^{2e}(\gamma_{4em}^2/8e)^{2e}\eta^{28e}|\mathrm{disc}(R)|^8|\mathrm{disc}(L)|^{1/m}\\
&\qquad\cdot C_{10}(2,e,m-2)\eta^{C_{11}(2,e,m-2)}(\gamma_{4em}^2/8e)^{4eC_{13}(2,e,m-2)}\\
&\qquad\cdot|\mathrm{disc}(R)|^{C_{12}(2,e,m-2)+2C_{13}(2,e,m-2)}|\mathrm{disc}(L)|^{(m+2)/m\cdot C_{13}(2,e,m-2)}.
\end{aligned}$$

Now we can calculate: for the multiplicative constant:

$$\begin{aligned}
C_{10}(2,e,m-2)(\gamma_e/8e)^{2e}(\gamma_{4em}^2/8e)^{e(6+4C_{13}(2,e,m-2))}&\\
=(2e(m-2)^2)^{e(m-2)m/2}(\gamma_e/8e)^{2e}\,(\gamma_{4em}^2/8e)^{2em}&\\
\leqslant(2em^2)^{e(m-2)m/2}\cdot 1\cdot(2em^2)^{2em}&\\
=(2em^2)^{e(m^2-2m+4m)/2}&\\
=C_{10}(2,e,m),&
\end{aligned}$$

for the exponent of η:

$$28e+C_{11}(2,e,m-2)=28e+14e(m-2)=C_{11}(2,e,m),$$

for the exponent of $|\mathrm{disc}(R)|$:

$$2+8+C_{12}(2,e,m-2)+2C_{13}(2,e,m-2)$$
$$=10+\frac{(m-2)(m+14)}{4}+(m-3)=\frac{m^2+16m}{4}=C_{12}(2,e,m),$$

for the exponent of $|\mathrm{disc}(L)|$:

$$\frac{2}{m}+\frac{1}{m}+\frac{m+2}{m}\cdot C_{13}(2,e,m-2)=\frac{3}{m}+\frac{(m+2)(m-3)}{2m}$$
$$=\frac{6+(m^2-m-6)}{2m}=\frac{(m-1)m}{2m}=C_{13}(2,e,m).$$

Case (c), part (iv). — For $i=j=1$ or 2, Lemma 4.4(iv)(c) gives

$$|\psi(v_i,v_i)|_D\leqslant 2^{-5/2}\gamma_e^{1/2}\gamma_{4em}^2\eta^7|\mathrm{disc}(R)|^{2/e}|\mathrm{disc}(L)|^{1/2em}. \tag{4.8}$$

This establishes (iv) for $i=j=1$ or 2 because, using Lemma 2.3 and the fact that $m\geqslant 2$ so $(m(m+1)+14)/8\geqslant 5/2$,

$$2^{-5/2}\gamma_e^{1/2}\gamma_{4em}^2\leqslant 2^{-5/2}e^{1/2}(4em)^2=2^{3/2}e^{5/2}m^2\leqslant(2em^2)^{5/2}\leqslant C_{14}(2,e,m),$$
$$7=C_{15}(2,e,m),$$
$$\frac{2}{e}=\frac{2\cdot 3+26}{16e}\leqslant\frac{m(m+1)+26}{16e}=C_{16}(2,e,m),$$
$$\frac{1}{2em}\leqslant\frac{2}{8e}\leqslant\frac{m+1}{8e}\leqslant C_{17}(d,e,m).$$

For $i=j\geqslant 3$, induction gives

$$|\psi(v_j,v_j)|_D\leqslant C_{14}(2,e,m-2)\eta^{C_{15}(2,e,m-2)}|\mathrm{disc}(R)|^{C_{16}(2,e,m-2)}$$
$$\cdot|\mathrm{disc}(M^\perp)|^{C_{17}(2,e,m-2)}$$
$$\leqslant C_{14}(2,e,m-2)\eta^{C_{15}(2,e,m-2)}|\mathrm{disc}(R)|^{C_{16}(2,e,m-2)}$$
$$\cdot\big((\gamma_{4em}^2/8e)^{4e}|\mathrm{disc}(R)|^2|\mathrm{disc}(L)|^{(m+2)/m}\big)^{C_{17}(2,e,m-2)}.$$

Now we can calculate: for the multiplicative constant:

$$C_{14}(2,e,m-2)(\gamma_{4em}^2/8e)^{4eC_{17}(2,e,m-2)}$$
$$=(2e(m-2)^2)^{((m-2)(m-1)+14)/8}\cdot(\gamma_{4em}^2/8e)^{(m-1)/2}$$
$$\leqslant(2em^2)^{(m^2-3m+16)/8}\cdot(2em^2)^{(4m-4)/8}$$
$$=(2em^2)^{(m^2+m+12)/8}$$
$$\leqslant C_{14}(2,e,m),$$

for the exponent of η:

$$C_{15}(2,e,m-2)=7=C_{15}(2,e,m),$$

for the exponent of $|\mathrm{disc}(R)|$:

$$C_{16}(2,e,m-2)+2C_{17}(2,e,m-2)=\frac{(m-2)(m-1)+26}{16e}+2\cdot\frac{m-1}{8e}$$
$$=\frac{m^2+m+24}{16e}\leqslant\frac{m(m+1)+26}{16e}=C_{16}(2,e,m),$$

for the exponent of $|\mathrm{disc}(L)|$:

$$\frac{m+2}{m}C_{17}(2,e,m-2)=\frac{m+2}{m}\cdot\frac{m-1}{8e}$$
$$=\frac{m^2+m-2}{8em}\leqslant\frac{m(m+1)}{8em}=C_{17}(2,e,m).\qquad\square$$

Lemma 4.6. — *Let $(D,\dagger)$ be a division $\mathbb{Q}$-algebra with a positive involution. Let V be a left D-vector space with a non-degenerate $(D,\dagger)$-skew-Hermitian form $\psi\colon V\times V\to D$. Let L be a $\mathbb{Z}$-lattice in V such that $\mathrm{Trd}_{D/\mathbb{Q}}\psi(L\times L)\subset\mathbb{Z}$. Let $R=\mathrm{Stab}_D(L)$. Then $\mathrm{disc}(L)R^\dagger\subset R$.*

Proof. — Let $a\in R$ and $x,y\in L$. Then

$$\mathrm{Trd}_{D/\mathbb{Q}}\psi(a^\dagger x,y)=\mathrm{Trd}_{D/\mathbb{Q}}\big(a^\dagger\psi(x,y)\big)=\mathrm{Trd}_{D/\mathbb{Q}}\big(\psi(x,y)a^\dagger\big)=\mathrm{Trd}_{D/\mathbb{Q}}\psi(x,ay).$$

Since $x,ay\in L$, we conclude that $\mathrm{Trd}_{D/\mathbb{Q}}\psi(a^\dagger x,y)\in\mathbb{Z}$.

Since this holds for all $y\in L$, we have $a^\dagger x\in L^*$. Consequently,

$$\mathrm{disc}(L)a^\dagger x=[L^*:L]a^\dagger x\in L.$$

This holds for all $x\in L$, so $\mathrm{disc}(L)a^\dagger\in\mathrm{Stab}_D(L)=R$. $\square$

To complete the proof of Theorem 1.2, we combine Proposition 4.5 and Lemma 4.6. The resulting exponent of $|\mathrm{disc}(L)|$ in (iii) is $C_{11}(d,e,m)+C_{13}(d,e,m)$ and the exponent of $|\mathrm{disc}(L)|$ in (iv) is $C_{15}(d,e,m)+C_{17}(d,e,m)$, while the other constants in Theorem 1.2 are the same as the corresponding constants in Proposition 4.5.

5. Application to the Zilber–Pink conjecture

In this section we study special subvarieties of PEL type from the point of view of Shimura data. The main result of the section is that Shimura datum components of simple PEL type I and II lie in a single $\mathbf{GSp}_{2g}(\mathbb{R})$-conjugacy class, which we describe explicitly. We also establish a bound on the dimension of all special subvarieties of PEL type in $\mathcal{A}_g$, demonstrating that Theorem 1.3 is indeed a consequence of the Zilber–Pink conjecture. We end the section by outlining the strategy of the proof of Theorem 1.3 carried out in the subsequent sections.

For our notation and terminology around Shimura datum components, see [DO21, sec. 2A and 2B].

5.A. Shimura data. — Let $L = \mathbb{Z}^{2g}$, let $V = L_{\mathbb{Q}}$ and let $\phi : L \times L \to \mathbb{Z}$ be the symplectic form represented, in the standard basis, by the matrix J_{2g}. Let $\mathbf{G} = \mathbf{GSp}(V, \phi) = \mathbf{GSp}_{2g}$ and let $\Gamma = \mathbf{Sp}_{2g}(\mathbb{Z})$. Let X^+ denote the $\mathbf{G}(\mathbb{R})^+$-conjugacy class of the morphism $h_0 \colon \mathbb{S} \to \mathbf{G}_{\mathbb{R}}$ given by

$$h_0(a+ib) \longmapsto \left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)^{\oplus g}. \tag{5.1}$$

Then $(\mathbf{G}, X^+)$ is a Shimura datum component and there is a $\mathbf{G}(\mathbb{R})^+$-equivariant bijection $X^+ \cong \mathcal{H}_g$, where $\mathcal{H}_g$ is the Siegel upper half-space. The moduli space of principally polarised abelian varieties of dimension g, denoted $\mathcal{A}_g$, is the Shimura variety whose complex points are $\Gamma\backslash X^+$.

Let S be a special subvariety of PEL type of $\mathcal{A}_g$, as defined in section 1.C, and let R be its generic endomorphism ring. Choose a point $x \in X^+$ whose image $s \in \mathcal{A}_g$ is an endomorphism generic point in $S(\mathbb{C})$. Then x induces an isomorphism $H_1(A_s, \mathbb{Z}) \cong L$ and hence the action of R on A_s induces an action of R on L.

Let $\mathbf{H}$ denote the centraliser in $\mathbf{G}$ of the action of R on L, which is a reductive $\mathbb{Q}$-algebraic group. We call $\mathbf{H}$ the *general Lefschetz group* of S. Note that $\mathbf{H}$ is only defined up to conjugation by Γ, because different choices of x may lead to isomorphisms $H_1(A_s, \mathbb{Z}) \cong L$ which differ by Γ. (The group $\mathbf{H}$ is isomorphic to the Lefschetz group of an endomorphism generic abelian variety parameterised by S, as defined in [Mil99], thanks to [Mil99, Th. 4.4]. However it seems to be more common to call $\mathbf{H} \cap \mathbf{Sp}$ or $(\mathbf{H} \cap \mathbf{Sp})^\circ$ the Lefschetz group, so we have added the adjective "general" by analogy with the general symplectic and general orthogonal groups.)

The special subvariety of PEL type S is a Shimura subvariety component of $\mathcal{A}_g$ associated with a Shimura subdatum component of the form $(\mathbf{H}^\circ, X^+_{\mathbf{H}}) \subset (\mathbf{G}, X^+)$, where $\mathbf{H}$ is the general Lefschetz group of S (see [Mil05, paragraph above Theorem 8.17]).

We say that $(\mathbf{H}, X^+_{\mathbf{H}}) \subset (\mathbf{G}, X^+)$ is a *Shimura subdatum component of simple PEL type I or II* if it is a Shimura subdatum component associated with a special subvariety of PEL type, where $\mathbf{H}$ is the general Lefschetz group, and its generic endomorphism algebra is a division algebra with positive involution of type I or II. Note that in the simple type I or II case, $\mathbf{H} = \mathbf{H}^\circ$.

5.B. Representatives of conjugacy classes of Shimura data of simple PEL type I or II. — The Shimura subdatum components of $(\mathbf{G}, X^+)$ of simple PEL type I or II lie in only finitely many $\mathbf{G}(\mathbb{R})^+$-conjugacy classes. Indeed, we shall now explicitly describe finitely many Shimura subdatum components which represent these $\mathbf{G}(\mathbb{R})^+$-conjugacy classes. Note that, for convenience, these representative subdatum components are not of simple PEL type, although they are of PEL type. This generalises [DO22, Lem. 6.1], which is the case $g = 2$, $d = 2$, $e = m = 1$.

Let d, e, m be positive integers such that $d^2em = 2g$, $d = 1$ or 2 and dm is even. For fixed g, there are only finitely many integers d, e, m satisfying these conditions. As we shall show, each triple d, e, m corresponds to a single $\mathbf{G}(\mathbb{R})^+$-conjugacy class of Shimura subdatum components of simple PEL type I or II.

Let $D_0 = \mathrm{M}_d(\mathbb{Q})^e$. Define a $\mathbb{Q}$-algebra homomorphism $\iota_0 \colon D_0 \to \mathrm{M}_{2g}(\mathbb{Q})$ as follows:
– when $d = 1$: $\iota_0(a_1, \dots, a_e) = a_1 I_m \oplus \cdots \oplus a_e I_m$.
– when $d = 2$:

$$\iota_0\Big(\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}, \dots, \begin{pmatrix} a_e & b_e \\ c_e & d_e \end{pmatrix}\Big) = \begin{pmatrix} a_1 I_{2m} & b_1 I_{2m} \\ c_1 I_{2m} & d_1 I_{2m} \end{pmatrix} \oplus \cdots \oplus \begin{pmatrix} a_e I_{2m} & b_e I_{2m} \\ c_e I_{2m} & d_e I_{2m} \end{pmatrix}.$$

We view V as a left D_0-module via ι_0.

Let t denote the involution of D_0 which is transpose on each factor. Since dm is even, $\iota_0(D_0)$ commutes with J_{2g} and so, for all $a \in D_0$ and $x, y \in V$, we have

$$\phi(ax, y) = x^t \iota(a)^t J_{2g} y = x^t J_{2g} \iota(a)^t y = \phi(x, a^t y).$$

Thus $\phi \colon V \times V \to \mathbb{Q}$ is a (D_0, t)-compatible symplectic form. By Lemma 3.1 and Corollary 3.3, there is a unique non-degenerate (D_0, t)-skew-Hermitian form $\psi_0 \colon V \times V \to D_0$ such that $\phi = \mathrm{Trd}_{D_0/\mathbb{Q}} \psi_0$.

Let $\mathbf{H}_0$ denote the centraliser of $\iota_0(D_0)$ in $\mathbf{G}$. In other words,

$$\mathbf{H}_0 = \{g_1^{\oplus d} \oplus g_2^{\oplus d} \oplus \cdots \oplus g_e^{\oplus d} : g_1, \dots, g_e \in \mathbf{GSp}_{dm},\ \nu(g_1) = \cdots = \nu(g_e)\}, \tag{5.2}$$

where $\nu \colon \mathbf{GSp}_{dm} \to \mathbb{G}_m$ denotes the symplectic multiplier character. This is a connected $\mathbb{Q}$-algebraic group, and it is equal to the general Lefschetz group of a special subvariety of PEL type in which endomorphism generic points correspond to abelian varieties isogenous to a product of the form $A_1^d \times \cdots \times A_e^d$ where $A_1, \dots, A_e$ are pairwise non-isogenous simple abelian varieties of dimension $dm/2$ with $\mathrm{End}(A_1) = \cdots = \mathrm{End}(A_e) = \mathbb{Z}$.

Lemma 5.1. — *Let $(\mathbf{H}, X_{\mathbf{H}}^+) \subset (\mathbf{GSp}_{2g}, \mathcal{H}_g)$ be a Shimura subdatum component of simple PEL type I or II. Let D be the generic endomorphism algebra of $(\mathbf{H}, X_{\mathbf{H}}^+)$ and let F be the centre of D. Then $\mathbf{H}_{\mathbb{R}}$ is a $\mathbf{G}(\mathbb{R})^+$-conjugate of the group $\mathbf{H}_0$ constructed above for the parameters*

$$d = \sqrt{\dim_F(D)} = 1 \text{ or } 2, \quad e = [F : \mathbb{Q}], \quad m = 2g/d^2 e.$$

Proof. — The tautological family of principally polarised abelian varieties on X^+ restricts to a family of principally polarised abelian varieties on $X_{\mathbf{H}}^+$. The polarisation induces a Rosati involution $\dagger$ of the endomorphism algebra of this family, namely D. As we saw in the construction of the general Lefschetz group, D acts on V. Via this action, the symplectic form $\phi \colon V \times V \to \mathbb{Q}$ is $(D, \dagger)$-compatible.

Since $(D, \dagger)$ is a simple $\mathbb{Q}$-algebra with a positive involution of type I or II, there is an isomorphism $\alpha \colon (D_{0,\mathbb{R}}, t) \to (D_{\mathbb{R}}, \dagger)$ of $\mathbb{R}$-algebras with involution (where $D_0 = \mathrm{M}_d(\mathbb{Q})^e$ for the parameters d and e specified in the lemma). We obtain an action of $D_{0,\mathbb{R}}$ on $V_{\mathbb{R}}$ by composing the action of $D_{\mathbb{R}}$ with α.

Since ϕ is $(D, \dagger)$-compatible, it is also $(D_{0,\mathbb{R}}, t)$-compatible under the via α. Hence there is a unique non-degenerate $(D_{0,\mathbb{R}}, t)$-skew-Hermitian form $\psi_\alpha \colon V_{\mathbb{R}} \times V_{\mathbb{R}} \to D_{0,\mathbb{R}}$ such that $\phi = \mathrm{Trd}_{D_{0,\mathbb{R}}/\mathbb{R}} \psi_\alpha$, where "$(D_{0,\mathbb{R}}, t)$-skew-Hermitian" refers to the action via α. (Note that ψ_α is in general different from ψ_0 because ψ_0 is $(D_{0,\mathbb{R}}, t)$-skew-Hermitian with respect to the action via ι_0.)

By Lemma 3.5, there exists a $D_{0,\mathbb{R}}$-basis $v_1, \dots, v_m$ for $V_{\mathbb{R}}$ with respect to the action via ι_0 which is symplectic (if $d = 1$) or unitary (if $d = 2$) with respect to ψ_0. There likewise exists a $D_{0,\mathbb{R}}$-basis $w_1, \dots, w_m$ for $V_{\mathbb{R}}$ with respect to the action via α which is symplectic or unitary with respect to ψ_α.

Define $\gamma \in \mathbf{GL}(V_{\mathbb{R}})$ by

$$\gamma(\iota_0(a_1)v_1 + \cdots + \iota_0(a_m)v_m) = \alpha(a_1)w_1 + \cdots + \alpha(a_m)w_m$$

for all $a_1, \dots, a_m \in D_{0,\mathbb{R}}$. Because $v_1, \dots, v_m$ and $w_1, \dots, w_m$ are symplectic or unitary bases (depending on d) with respect to ψ_0 and ψ_α respectively, we have

$$\psi_\alpha(\gamma(v_i), \gamma(v_j)) = \psi_\alpha(w_i, w_j) = \psi_0(v_j, v_j)$$

for all i, j. Because ψ_0 and ψ_α are $(D_{0,\mathbb{R}}, t)$-skew-Hermitian with respect to the actions via ι_0 and α respectively, it follows that

$$\psi_\alpha(\gamma(v), \gamma(w)) = \psi_0(v, w)$$

for all $v, w \in V_{\mathbb{R}}$. Taking the reduced trace, we obtain $\phi(\gamma(v), \gamma(w)) = \phi(v, w)$ for all $v, w \in V_{\mathbb{R}}$. In other words, $\gamma \in \mathbf{Sp}(V_{\mathbb{R}}, \phi) \subset \mathbf{G}(\mathbb{R})^+$.

Since γ is an isomorphism between the representations of $D_{0,\mathbb{R}}$ given by α and ι_0, $\gamma\mathbf{H}_0\gamma^{-1}$ is the centraliser in $\mathbf{G}$ of the action of $D_{0,\mathbb{R}}$ via α. In other words, $\gamma\mathbf{H}_0\gamma^{-1}$ is the centraliser in $\mathbf{G}$ of the action of $D_{\mathbb{R}}$, which is the general Lefschetz group $\mathbf{H}$. □

Lemma 5.2. — *For each triple of positive integers d, e, m satisfying $d^2em = 2g$, $d = 1$ or 2 and dm even, there exists a unique Shimura subdatum component $(\mathbf{H}_0, X_0^+)$ of $(\mathbf{G}, X^+)$ with group $\mathbf{H}_0$. Furthermore, the Hodge parameter h_0 from (5.1) is in X_0^+.*

Proof. — First note that $h_0 \in X^+$ and h_0 factors through $\mathbf{H}_{0,\mathbb{R}}$. Hence if X_0^+ denotes the $\mathbf{H}_0(\mathbb{R})^+$-conjugacy class of h_0 in $\mathrm{Hom}(\mathbb{S}, \mathbf{H}_{0,\mathbb{R}})$, then $(\mathbf{H}_0, X_0^+)$ is a Shimura subdatum component of $(\mathbf{G}, X^+)$.

To establish the uniqueness, let X_0^+ now denote any subset of X^+ such that $(\mathbf{H}_0, X_0^+)$ is a Shimura datum component. Let $\mathbf{H}_0^{\mathrm{ad}}$ denote the quotient of $\mathbf{H}_0$ by its centre. By [Mil05, Prop. 5.7(a)], X_0^+ is in bijection with its image $(X_0^+)^{\mathrm{ad}} \subset \mathrm{Hom}(\mathbb{S}, \mathbf{H}_{0,\mathbb{R}}^{\mathrm{ad}})$ under composition with the natural map $\mathbf{H}_{0,\mathbb{R}} \to \mathbf{H}_{0,\mathbb{R}}^{\mathrm{ad}}$.

Observe that $\mathbf{H}_0^{\mathrm{ad}} \cong \mathbf{PGSp}_{md}^e$. Therefore, $(X_0^+)^{\mathrm{ad}}$ is a product of $\mathbf{PGSp}_{md}(\mathbb{R})^+$-conjugacy classes of morphisms $\mathbb{S} \to \mathbf{PGSp}_{md,\mathbb{R}}$ satisfying conditions SV1–SV3 from [Mil05, §4]. From [Mil05, Prop 1.24], and the following paragraphs, there exists only one $\mathbf{PGSp}_{md}(\mathbb{R})$-conjugacy class X_{md} of morphisms $\mathbb{S} \to \mathbf{PGSp}_{md,\mathbb{R}}$ satisfying SV1–SV3. It has two connected components X_{md}^+ and X_{md}^- corresponding to the connected components of $\mathbf{PGSp}_{md}(\mathbb{R})$. In other words, $(X_0^+)^{\mathrm{ad}}$ is equal to a direct product of copies of the spaces X_{md}^+ and X_{md}^-.

Consider the morphisms $h_2^+, h_2^- : \mathbb{S} \to \mathbf{GL}_{2,\mathbb{R}}$ defined by

$$h_2^+ : a + ib \longmapsto \left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right) \text{ and } h_2^- : a + ib \longmapsto \left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right).$$

Then $(h_2^+)^{\oplus md/2}$ and $(h_2^-)^{\oplus md/2}$ are non-$\mathbf{GSp}_{md}(\mathbb{R})^+$-conjugate morphisms $\mathbb{S} \to \mathbf{GSp}_{md,\mathbb{R}}$ satisfying SV1–SV3. Therefore, the images of their $\mathbf{GSp}_{md}(\mathbb{R})^+$-conjugacy classes in $\mathrm{Hom}(\mathbb{S}, \mathbf{PGSp}_{md,\mathbb{R}})$ are precisely X_{md}^+ and X_{md}^-.

It follows that $(X_0^+)^{\mathrm{ad}}$ is the image in $\mathrm{Hom}(\mathbb{S}, \mathbf{PGSp}^e_{md,\mathbb{R}})$ of the $\mathbf{GSp}^e_{md}(\mathbb{R})^+$-conjugacy class of an element $h \in \mathrm{Hom}(\mathbb{S}, \mathbf{GSp}^e_{md,\mathbb{R}})$ of the form

$$\big((h_2^{\pm})^{\oplus md/2}, \dots, (h_2^{\pm})^{\oplus md/2}\big),$$

for some sequence of signs in $\{\pm\}^e$. Since the image of h in $\mathrm{Hom}(\mathbb{S}, \mathbf{GSp}_{md^2e,\mathbb{R}})$ (obtained by repeating each component d times block-diagonally) lies in a Shimura datum, it satisfies condition SV2 of [Mil05], that is, the stabiliser of h in $\mathbf{GSp}_{md^2e}(\mathbb{R})$ is compact modulo the centre. This only holds when

$$h = h^+ =: ((h_2^+)^{\oplus md/2}, \dots, (h_2^+)^{\oplus md/2}) \text{ or } h = h^- := ((h_2^-)^{\oplus md/2}, \dots, (h_2^-)^{\oplus md/2}).$$

This can be checked by observing that the centraliser of $h_2^+ \oplus h_2^-$ in $\mathbf{GSp}_4(\mathbb{R})$ is non-compact modulo the centre.

Note that the image of h^+ in $\mathrm{Hom}(\mathbb{S}, \mathbf{GSp}_{md^2e,\mathbb{R}})$ is equal to h_0. Since $h_0 \in X^+$ while the image of h^- is not in X^+, we conclude that X_0^+ must be equal to the $\mathbf{H}_0(\mathbb{R})$-conjugacy class of h_0. □

Corollary 5.3. — *If $(\mathbf{H}, X_{\mathbf{H}}^+) \subset (\mathbf{G}, X^+)$ is a Shimura subdatum component of simple PEL type I or II and $\mathbf{H} = g\mathbf{H}_0 g^{-1}$ for $g \in \mathbf{G}(\mathbb{R})^+$, then $X_{\mathbf{H}}^+ = gX_0^+$ where $(\mathbf{H}_0, X_0^+) \subset (\mathbf{G}, X^+)$ is the unique Shimura subdatum component given by Lemma 5.2.*

5.C. Dimension of special subvarieties of PEL type. — In this section we prove Proposition 1.7: Proposition 5.4 is Proposition 1.7(i), while Proposition 5.5 is Proposition 1.7(ii).

Proposition 5.4. — *Let $S \subset \mathcal{A}_g$ be a special subvariety, not equal to $\mathcal{A}_g$. If S is of simple PEL type, then* $\dim(S) \leqslant \dim(\mathcal{A}_g) - g^2/4$.

Proof. — Let D be the generic endomorphism algebra of S. Following our usual notation, let F be the centre of D and let

$$d = \sqrt{\dim_F(D)}, \quad e = [F : \mathbb{Q}], \quad m = 2g/d^2e.$$

When D has Albert type IV, we need some additional notation. Let $s \in S(\mathbb{C})$. Then $D_{\mathbb{R}} \cong \mathrm{M}_d(\mathbb{C})^{e/2}$ acts $\mathbb{R}$-linearly on the tangent space $T_0(A_s(\mathbb{C}))$. For each $i = 1, \dots, e/2$, let r_i denote the multiplicity in $T_0(A_s(\mathbb{C}))$ of the standard representation of the i-th factor $\mathrm{M}_d(\mathbb{C})$ of $D_{\mathbb{R}}$. Similarly let s_i denote the multiplicity of the complex conjugate of the standard representation of the i-th factor of $D_{\mathbb{R}}$. The values r_i and s_i are independent of the choice of $s \in S(\mathbb{C})$, and satisfy $r_i + s_i = dm$.

The dimension of special subvarieties of simple PEL type was determined by Shimura [Shi63, 4.1]. Note that our m is the same as m in [Shi63], while our e is called g in [Shi63]. For a more modern account of this theory, see [BL04, Chap. 9]. For each type of endomorphism algebra D, we quote the dimension of the special subvariety from [Shi63, 4.1] and use some elementary inequalities.

When D has type I, $d = 1$, $em = 2g$ and $e \geqslant 2$ since $S \neq \mathcal{A}_g$, so $m \leqslant g$. Hence

$$\dim(S) = \tfrac{1}{2}\tfrac{m}{2}\big(\tfrac{m}{2} + 1\big)e \leqslant \tfrac{1}{2}g\big(\tfrac{1}{2}g + 1\big) = \tfrac{1}{4}g^2 + \tfrac{1}{2}g.$$

When D has type II, $d = 2$, $em = g/2$ and $m \leqslant g/2$ so

$$\dim(S) = \tfrac{1}{2}m(m+1)e \leqslant \tfrac{1}{4}g\big(\tfrac{1}{2}g+1\big) = \tfrac{1}{8}g^2 + \tfrac{1}{4}g.$$

When D has type III, $d = 2$, $em = g/2$ and $m \leqslant g/2$ so

$$\dim(S) = \tfrac{1}{2}m(m-1)e \leqslant \tfrac{1}{4}g\big(\tfrac{1}{2}g-1\big) = \tfrac{1}{8}g^2 - \tfrac{1}{4}g.$$

When D has type IV, $2g = d^2em$ and $e \geqslant 2$ since F is a CM field, so $m \leqslant g$. Furthermore $r_i + s_i = dm$ so $r_i s_i \leqslant d^2m^2/4$ for each i. Hence,

$$\dim(S) = \sum_{i=1}^{e/2} r_i s_i \leqslant \tfrac{1}{2}e \cdot \tfrac{1}{4}d^2m^2 = \tfrac{1}{4}gm \leqslant \tfrac{1}{4}g^2.$$

Hence in all cases,

$$\dim(S) \leqslant \tfrac{1}{4}g^2 + \tfrac{1}{2}g = \tfrac{1}{2}g(g+1) - \tfrac{1}{4}g^2 = \dim(\mathcal{A}_g) - \tfrac{1}{4}g^2. \qquad \square$$

Proposition 5.5. — *Let $S \subset \mathcal{A}_g$ be a special subvariety, not equal to $\mathcal{A}_g$. If S is of PEL type, then* $\dim(S) \leqslant \dim(\mathcal{A}_g) - g + 1$.

Proof. — Note that $g^2/4 \geqslant g - 1$ for all real numbers g, so Proposition 5.4 implies the claim for special subvarieties of simple PEL type.

Let $S \subset \mathcal{A}_g$ be a special subvariety of non-simple PEL type. By adding level structure, we may obtain a finite cover $S' \to S$ which is a fine moduli space of abelian varieties with PEL structure. Then there is a universal abelian scheme with PEL structure $A \to S'$. Since S' is of non-simple PEL type, A is a non-simple abelian scheme. Thus there exist non-trivial abelian schemes $A_1, A_2 \to S'$ such that A is isogenous to $A_1 \times A_2$. (There may be multiple choices of isogeny decompositions of A. Choose any such decomposition.) Let g_1, g_2 denote the relative dimensions of A_1 and A_2 respectively.

Set

$$T = \{(s, s_1, s_2) \in S' \times \mathcal{A}_{g_1} \times \mathcal{A}_{g_2} : A_s \text{ is isogenous to } A_{s_1} \times A_{s_2}\}.$$

Since isogenies $A_s \to A_{s_1} \times A_{s_2}$ give rise to Hodge classes on $A_s \times A_{s_1} \times A_{s_2}$, the locus T is a countable union of special subvarieties of $S' \times \mathcal{A}_{g_1} \times \mathcal{A}_{g_2}$.

By construction, the projection $T \to S'$ is surjective on $\mathbb{C}$-points. An irreducible complex algebraic variety cannot be contained in the union of countably many proper closed subvarieties. Hence there exists an irreducible component $T^+ \subset T$ such that the image of T^+ is dense in S'. Hence $\dim(T^+) \geqslant \dim(S') = \dim(S)$.

Given any two abelian varieties A_{s_1} and A_{s_2} over $\mathbb{C}$, there are only countably many isomorphism classes of abelian varieties which are isogenous to $A_{s_1} \times A_{s_2}$. Furthermore each abelian variety of dimension g carries only finitely many PEL structures parameterised by S' (the natural morphism $S' \to \mathcal{A}_g$ is finite). Hence the projection $T \to \mathcal{A}_{g_1} \times \mathcal{A}_{g_2}$ has countable fibres. Therefore

$$\dim(T^+) \leqslant \dim(\mathcal{A}_{g_1} \times \mathcal{A}_{g_2}) = \frac{g_1(g_1+1)}{2} + \frac{g_2(g_2+1)}{2}.$$

Since $g_1 + g_2 = g$, we obtain

$$\tfrac{1}{2}g_1(g_1+1) + \tfrac{1}{2}g_2(g_2+1) = \tfrac{1}{2}\big((g_1+g_2)^2 - 2g_1g_2 + (g_1+g_2)\big) = \tfrac{1}{2}(g^2+g) - g_1g_2.$$

Therefore

$$\dim(S) \leqslant \dim(T^+) \leqslant \dim(\mathcal{A}_g) - g_1g_2.$$

Now $g_1g_2 = g_1(g-g_1)$ is a quadratic function of g_1 with a maximum at $g_1 = g/2$. Hence, for $1 \leqslant g_1 \leqslant g-1$, g_1g_2 is minimised when $g_1 = 1$ or $g-1$. Thus $g_1g_2 \geqslant g-1$. □

6. Construction of representation and closed orbit

This section constructs the representation required for the strategy outlined in section 1.E and proves that it satisfies conditions (i) and (ii) of [DO22, Th. 1.2]. These conditions are algebraic and geometric in nature. We also prove a small piece of arithmetic information about the representation, namely Proposition 6.1(v), which will be used to obtain more substantial arithmetic properties in section 7. This section generalises [DO22, §§5.2 & 5.3].

We will actually construct two representations $\rho_L, \rho_R \colon \mathbf{G} \to \mathbf{GL}(W)$, which are induced by left and right multiplication respectively in $\mathrm{End}(V)$. The representation to which we shall apply [DO22, Th. 1.2] is ρ_L, while ρ_R is an auxiliary object required at the end of section 7.

Proposition 6.1. — *Let d, e and m be positive integers such that dm is even. Let $n = d^2em$. Let $L = \mathbb{Z}^n$ and let $\phi \colon L \times L \to \mathbb{Z}$ be the standard symplectic form as in section* 5.A. *Let $V = L_{\mathbb{Q}}$, let $\mathbf{G} = \mathbf{GSp}(V,\phi) = \mathbf{GSp}_{n,\mathbb{Q}}$ and let $\Gamma = \mathbf{Sp}_n(\mathbb{Z})$.*

Let E_0 be a $\mathbb{Q}$-subalgebra of $\mathrm{End}(V) = \mathrm{M}_n(\mathbb{Q})$ such that $E_{0,\mathbb{C}} \cong \mathrm{M}_{dm}(\mathbb{C})^e$ and the resulting $E_{0,\mathbb{C}}$-module structure on $V_{\mathbb{C}}$ is isomorphic to the direct sum of d copies of each of the e irreducible representations of $E_{0,\mathbb{C}}$. Let $\mathbf{H}_0$ be the $\mathbb{Q}$-algebraic subgroup of $\mathbf{G}$ whose k-points are

$$\mathbf{H}_0(k) = (E_0 \otimes_{\mathbb{Q}} k) \cap \mathbf{G}(k)$$

for each field extension k of $\mathbb{Q}$.

Then there exists a $\mathbb{Q}$-vector space W, a $\mathbb{Z}$-lattice $\Lambda \subset W$, $\mathbb{Q}$-algebraic representations $\rho_L, \rho_R \colon \mathbf{G} \to \mathbf{GL}(W)$, a vector $w_0 \in \Lambda$ and a constant C_{18} such that:

(i) $\mathrm{Stab}_{\mathbf{G},\rho_L}(w_0) = \mathrm{Stab}_{\mathbf{G},\rho_R}(w_0) = \mathbf{H}_0$;

(ii) *the orbit $\rho_L(\mathbf{G}(\mathbb{R}))w_0$ is closed in $W_{\mathbb{R}}$;*

(iii) *ρ_L and ρ_R commute with each other;*

(iv) *$\rho_L(\Gamma)$ and $\rho_R(\Gamma)$ stabilise Λ;*

(v) *for each $u \in \mathbf{G}(\mathbb{R})$, if the group $\mathbf{H}_u = u\mathbf{H}_{0,\mathbb{R}}u^{-1}$ is defined over $\mathbb{Q}$, then there exists $d_u \in \mathbb{R}_{>0}$ such that*

$$d_u\rho_L(u)\rho_R(u)w_0 \in \Lambda \quad \text{and} \quad d_u \leqslant C_{18}|\mathrm{disc}(S_u)|^{1/2},$$

where S_u denotes the ring $uE_{0,\mathbb{R}}u^{-1} \cap \mathrm{M}_n(\mathbb{Z})$.

In our application to Theorem 1.3, $\mathbf{H}_0$ shall be equal to the group $\mathbf{H}_0$ defined in (5.2). To achieve this, let $d = 1$ or 2 and define D_0 and $\iota_0 \colon D_0 \to \mathrm{M}_n(\mathbb{Q})$ as in section 5.B (with $n = 2g$). Let E_0 be the centraliser of $\iota_0(D_0)$ in $\mathrm{M}_n(\mathbb{Q})$, that is,

$$E_0 = \{f_1^{\oplus d} \oplus f_2^{\oplus d} \oplus \cdots \oplus f_e^{\oplus d} \in \mathrm{M}_n(\mathbb{Q}) : f_1, \ldots, f_e \in \mathrm{M}_{dm}(\mathbb{Q})\}. \tag{6.1}$$

It is immediate that intersecting this algebra E_0 with $\mathbf{G}$ yields the same group $\mathbf{H}_0$ as in (5.2). Furthermore, the map $(f_1, \ldots, f_e) \mapsto f_1^{\oplus d} \oplus f_2^{\oplus d} \oplus \cdots \oplus f_e^{\oplus d}$ is an isomorphism of $\mathbb{Q}$-algebras $\mathrm{M}_{dm}(\mathbb{Q})^e \to E_0$. By decomposing V as a direct sum of dm-dimensional subspaces, matching the block diagonal decomposition of elements of E_0, we see that V is isomorphic to the sum of d copies of each of the e irreducible representations of E_0. After extending scalars to $\mathbb{C}$, we conclude that E_0 as defined by (6.1) satisfies the conditions of Proposition 6.1.

Allowing more general choices of E_0 in Proposition 6.1 than simply (6.1), and only imposing conditions on E_0 after extending scalars to $\mathbb{C}$, ensures that the proposition could be used as part of a similar strategy for proving the Zilber–Pink conjecture for special subvarieties of simple PEL type III and IV, as well as types I and II.

6.A. Construction of the representation. — Let $\sigma_L, \sigma_R \colon \mathbf{G} \to \mathbf{GL}(\mathrm{End}(V))$ denote the left and right multiplication representations of $\mathbf{G}$:

$$\sigma_L(g)f = gf, \quad \sigma_R(g)f = fg^{-1}.$$

Note that $\sigma_R(g)f = fg^{-1}$ rather than fg so that σ_R is a group representation. The representations ρ_L and ρ_R in Proposition 6.1 are induced by σ_L and σ_R via a linear algebra construction which we shall now explain, and hence one may think of $\rho_L(u)\rho_R(u)$ in Proposition 6.1(v) as being induced by conjugation by $u \in \mathbf{G}(\mathbb{R})$.

Let $\nu \colon \mathbf{G} = \mathbf{GSp}_n \to \mathbb{G}_m$ denote the symplectic multiplier character. Let $W = \bigwedge^{mn} \mathrm{End}(V)$, which is a $\mathbb{Q}$-vector space of dimension $\binom{n^2}{mn}$. The representations required by Proposition 6.1 are defined as

$$\rho_L = \bigwedge^{mn} \sigma_L \otimes \nu^{-mn/2}, \ \rho_R = \bigwedge^{mn} \sigma_R \otimes \nu^{mn/2} : \mathbf{G} \longrightarrow \mathbf{GL}(W).$$

The powers of ν are chosen so that both ρ_L and ρ_R restrict to the trivial representation on the scalars $\mathbb{G}_m \subset \mathbf{GSp}_n$.

Next we construct a vector $w_0 \in W$ satisfying Proposition 6.1(i). Observe that $\dim_{\mathbb{Q}}(E_0) = e(dm)^2 = mn$ so $\bigwedge^{mn} E_0$ is a 1-dimensional subspace of W. This was the reason we used the mn-th exterior power to define W.

Because E_0 is a subring of $\mathrm{End}(V)$, for any field extension k of $\mathbb{Q}$,

$$\mathrm{Stab}_{\mathbf{G}(k),\sigma_L}(E_0) = \mathbf{G}(k) \cap (E_0 \otimes_{\mathbb{Q}} k) = \mathbf{H}_0(k).$$

Similarly $\mathrm{Stab}_{\mathbf{G}(k),\sigma_R}(E_0) = \mathbf{H}_{d,e,m}(k)$. Consequently,

$$\mathrm{Stab}_{\mathbf{G},\rho_L}(\textstyle\bigwedge^{mn} E_0) = \mathrm{Stab}_{\mathbf{G},\rho_R}(\bigwedge^{mn} E_0) = \mathbf{H}_0.$$

The action of E_0 on $\bigwedge^{mn} E_0$ via the mn-th exterior power of the left regular representation is multiplication by the non-reduced norm $\mathrm{Nm}_{E_0/\mathbb{Q}}$. Choose an isomorphism

$\eta\colon E_{0,\mathbb{C}} \to \mathrm{M}_{dm}(\mathbb{C})$. Let $f \in E_{0,\mathbb{C}}$ and $\eta(f) = (f_1, \dots, f_e) \in \mathrm{M}_{dm}(\mathbb{C})^e$. Since the irreducible representations of $E_{0,\mathbb{C}}$ are projections onto the simple factors of $\mathrm{M}_{dm}(\mathbb{C})^e$, and each irreducible representation appears d times in $V_{\mathbb{C}}$, we have

$$\det(f) = \prod_{i=1}^{e} \det(f_i)^d.$$

Hence

$$\mathrm{Nm}_{E_{0,\mathbb{C}}/\mathbb{C}}(f) = \prod_{i=1}^{e} \mathrm{Nm}_{\mathrm{M}_{dm}(\mathbb{C})/\mathbb{C}}(f_i) = \prod_{i=1}^{e} \det(f_i)^{dm} = \det(f)^m.$$

If $f \in \mathbf{H}_0(\mathbb{Q}) \subset \mathbf{G}(\mathbb{Q})$, then $\det(f) = \nu(f)^{n/2}$ so

$$\mathrm{Nm}_{E_0/\mathbb{Q}}(f) = \nu(f)^{mn/2}.$$

Hence the action of $\mathbf{H}_0$ on $\bigwedge^{mn} E_0$ via ρ_L is multiplication by $\mathrm{Nm}_{E_0/\mathbb{Q}} \otimes \nu^{-mn/2} = 1$. Thus for any $w \in \bigwedge^{mn} E_0$, we have $\rho_L(\mathbf{H}_0)w = w$, while

$$\mathrm{Stab}_{\mathbf{G},\rho_L}(w) \subset \mathrm{Stab}_{\mathbf{G},\rho_L}(\textstyle\bigwedge^{mn} E_0) = \mathbf{H}_0.$$

Thus $\mathrm{Stab}_{\mathbf{G},\rho_L}(w) = \mathbf{H}_0$.

For similar reasons, the action of $\mathbf{H}_0$ on $\bigwedge^{mn} E_0$ via $\bigwedge^{mn} \sigma_R$ is multiplication by $\mathrm{Nm}_{E_0/\mathbb{Q}}^{-1}$, and hence the action of $\mathbf{H}_0$ on $\bigwedge^{mn} E_0$ via ρ_R is trivial. It follows that for any $w \in \bigwedge^{mn} E_0$, $\mathrm{Stab}_{\mathbf{G},\rho_R}(w) = \mathbf{H}_0$.

Let $\Lambda = \bigwedge^{mn} \mathrm{M}_n(\mathbb{Z})$, which is a $\mathbb{Z}$-lattice in W. Let $S_0 = E_0 \cap \mathrm{M}_n(\mathbb{Z})$, which is an order in E_0. Then $\bigwedge^{mn} S_0$ is a free $\mathbb{Z}$-module of rank 1 contained in Λ. Choose w_0 to be a generator of $\bigwedge^{mn} S_0$ (it does not matter which generator we choose).

Since $w_0 \in \bigwedge^{mn} E_0$, the argument above shows that w_0 satisfies Proposition 6.1(i). It is clear that ρ_L and ρ_R commute, so Proposition 6.1(iii) holds. It is also immediate that Proposition 6.1(iv) holds. Most of this section will be devoted to proving Proposition 6.1(ii). Since the proof of 6.1(v) is short, let us first include it here.

Proof of Proposition 6.1(v). — By definition,

$$\rho_L(u)\rho_R(u) = \textstyle\bigwedge^{mn} \sigma_L(u)\sigma_R(u) \in \mathbf{GL}(\bigwedge^{mn} \mathrm{End}(V)),$$

where $\sigma_L(u)\sigma_R(u) \in \mathbf{GL}(\mathrm{End}(V))$ is conjugation by u. Hence $\rho_L(u)\rho_R(u)w_0$ is a generator of the $\mathbb{Z}$-module $\bigwedge^{mn} uS_0u^{-1}$.

Let $d_u = \mathrm{covol}(S_u)/\mathrm{covol}(uS_0u^{-1})$ with respect to the volume form induced by the non-reduced trace form on $S_{u,\mathbb{R}}$. Then $d_u\rho_L(u)\rho_R(u)w_0$ is a generator for $\bigwedge^{mn} S_u$ and therefore is in Λ.

Conjugation by u pulls back $\mathrm{Tr}_{S_{u,\mathbb{R}}/\mathbb{R}}$ to $\mathrm{Tr}_{S_{0,\mathbb{R}}/\mathbb{R}}$. Hence

$$d_u = \mathrm{covol}(S, \mathrm{Tr}_{S_{u,\mathbb{R}}/\mathbb{R}})/\mathrm{covol}(S_0, \mathrm{Tr}_{S_{0,\mathbb{R}}/\mathbb{R}}) = \sqrt{|\mathrm{disc}(S_u)|/|\mathrm{disc}(S_0)|}. \qquad \square$$

6.B. Proof of closed orbit. — According to [BHC62, Prop. 6.3], in order to show that $\rho_L(\mathbf{G}(\mathbb{R}))w_0$ is closed in $W_{\mathbb{R}}$ (in the real topology), it suffices to prove that $\rho_L(\mathbf{G}(\mathbb{C}))w_0$ is Zariski closed in $W_{\mathbb{C}}$. Therefore, for the rest of this section, we shall deal entirely with linear algebra and algebraic geometry over $\mathbb{C}$.

Set

$$Q = \{g \in \mathrm{End}(V_{\mathbb{C}}) : \exists s \in \mathbb{C} \text{ s.t. for all } v, v' \in V_{\mathbb{C}}, \phi(gv, gv') = s\phi(v, v')\}.$$

Note that Q is equal to the union of $\mathbf{G}(\mathbb{C})$ with the set of elements of $\mathrm{End}(V_{\mathbb{C}})$ whose image is contained in a ϕ-isotropic subspace of $V_{\mathbb{C}}$. In particular,

$$\mathbf{G}(\mathbb{C}) = \{g \in Q : \det(g) \neq 0\}.$$

Let $e_1, \dots, e_n$ be a symplectic basis for $(V_{\mathbb{C}}, \phi)$. Then Q is a Zariski closed subset of $\mathrm{End}(V_{\mathbb{C}})$ because it is defined by the polynomial equations

$$\begin{aligned} &\phi(ge_i, ge_j) = 0 \quad \text{for each } i, j \text{ except when } \{i, j\} = \{2k-1, 2k\} \text{ for some } k, \\ &\phi(ge_1, ge_2) = \phi(ge_3, ge_4) = \cdots = \phi(ge_{n-1}, ge_n). \end{aligned}$$

Furthermore, Q is a homogeneous subset of $\mathrm{End}(V_{\mathbb{C}})$, that is, it is closed under multiplication by scalars.

Consequently, for any map from $\mathrm{End}(V_{\mathbb{C}})$ to another vector space whose coordinates are given by homogeneous polynomials of the same positive degree, the image of Q is homogeneous and Zariski closed. (This is because such a map induces a morphism of varieties between the associated projective spaces, and the image of the projective algebraic set $(Q \smallsetminus \{0\})/\mathbb{G}_m$ under such a morphism will again be a projective algebraic set.)

Note that $\sigma_L : \mathbf{G}(\mathbb{C}) \to \mathbf{GL}(\mathrm{End}(V_{\mathbb{C}}))$ extends to a $\mathbb{C}$-algebra homomorphism $\mathrm{End}(V_{\mathbb{C}}) \to \mathrm{End}(\mathrm{End}(V_{\mathbb{C}})) \cong \mathrm{M}_{n^2}(\mathbb{C})$ defined by the formula $\sigma_L(g)f = gf$. Considering σ_L as a representation of the multiplicative monoid $\mathrm{End}(V_{\mathbb{C}})$, it induces a monoid representation

$$\textstyle\bigwedge^{mn}\sigma_L \colon \mathrm{End}(V_{\mathbb{C}}) \longrightarrow \mathrm{End}(\bigwedge^{mn} \mathrm{End}(V_{\mathbb{C}})).$$

Here $\bigwedge^{mn}\sigma_L$ is a homogeneous morphism of degree mn, so the set $(\bigwedge^{mn}\sigma_L)(Q)w_0$ is a homogeneous Zariski closed subset of $W_{\mathbb{C}}$.

Lemma 6.2. — *There exist vectors $u_1, \dots, u_m \in V$ such that the map $\delta \colon \mathrm{End}(V) \to V^m$ defined by $\delta(f) = (f(u_1), \dots, f(u_m))$ restricts to an isomorphism of $\mathbb{Q}$-vector spaces $E_0 \to V^m$.*

Proof. — By the hypothesis of Proposition 6.1, we can decompose $V_{\mathbb{C}}$ as a direct sum of irreducible $E_{0,\mathbb{C}}$-modules

$$V_{\mathbb{C}} = \bigoplus_{i=1}^{e} \bigoplus_{j=1}^{d} V_{ij} \tag{6.2}$$

such that the action of $E_{0,\mathbb{C}} \cong \mathrm{M}_{dm}(\mathbb{C})^e$ on V_{ij} factors through the i-th copy of $\mathrm{M}_{dm}(\mathbb{C})$. Since $\mathrm{M}_{dm}(\mathbb{C})$ is a simple algebra, it has a unique irreducible representation (up to isomorphism), so we may choose an isomorphism of $\mathrm{M}_{dm}(\mathbb{C})$-modules $\theta_{ij} \colon \mathbb{C}^{dm} \to V_{ij}$.

Label the standard basis of $\mathbb{C}^{dm}$ as $e_{k\ell}$ for $1 \leqslant k \leqslant d$, $1 \leqslant \ell \leqslant m$. Given $f \in E_{0,\mathbb{C}}$, write $\eta(f) = (f_1, \dots, f_e) \in \mathrm{M}_{dm}(\mathbb{C})^e$. For $i = 1, \dots, e$, $k = 1, \dots, d$ and $\ell = 1, \dots, m$, let $f_{i,k\ell} \in \mathbb{C}^{dm}$ denote the column of f_i indexed by k and ℓ (ordered to match the

basis vectors $e_{k\ell}$). The action of $E_{0,\mathbb{C}}$ on V_{ij} factors through the i-th copy of $\mathrm{M}_{dm}(\mathbb{C})$ and θ_{ij} is an $\mathrm{M}_{dm}(\mathbb{C})$-module homomorphism, so

$$f(\theta_{ij}(e_{k\ell})) = \theta_{ij}(f_i(e_{k\ell})) = \theta_{ij}(f_{i,k\ell}).$$

For $\ell = 1, \dots, m$, let

$$u_\ell = \sum_{i=1}^{e} \sum_{j=1}^{d} \theta_{ij}(e_{j\ell}) \in V_{\mathbb{C}}.$$

(Note that the index j is used twice in this expression.) Then

$$f(u_\ell) = \sum_{i=1}^{e} \sum_{j=1}^{d} \theta_{ij}(f_{i,j\ell}). \tag{6.3}$$

If $f \in \ker(\delta) \cap E_{0,\mathbb{C}}$, then $f(u_\ell) = 0$ for $\ell = 1, \dots, m$. Since (6.2) is a direct sum and the θ_{ij} are injective, it follows from (6.3) that $f_{i,j\ell} = 0$ for all i, j and ℓ. In other words $f = 0$.

Thus $\delta|_{E_{0,\mathbb{C}}}$ is injective. In particular $\delta|_{E_0}$ is injective. Since

$$\dim_{\mathbb{Q}}(E_0) = \dim_{\mathbb{C}}(E_{0,\mathbb{C}}) = ed^2m^2 = \dim_{\mathbb{Q}}(V^m)$$

and δ is a linear map, it follows that $\delta|_{E_0}$ is an isomorphism $E_0 \to V^m$. □

Lemma 6.3. — *There exists a linear function $\zeta\colon W \to \mathbb{Q}$ such that $\zeta(w_0) \neq 0$ and*

$$\zeta\big((\Lambda^{mn}\sigma_L)(g)w\big) = \det(g)^m \zeta(w)$$

for all $g \in \mathrm{End}(V)$ and all $w \in W$.

Proof. — Define ζ to be the linear map on mn-th exterior powers induced by δ from Lemma 6.2. Then ζ is a linear map $W = \Lambda^{mn}\mathrm{End}(V) \to \Lambda^{mn}V^m \cong \mathbb{Q}$. We identify $\Lambda^{mn}V^m$ with $\mathbb{Q}$ (the choice of isomorphism $\Lambda^{mn}V^m \cong \mathbb{Q}$ is not important).

Since $\delta|_{E_0}$ is an isomorphism $E_0 \to V^m$ and w_0 is a generator of $\Lambda^{mn}E_0$, we deduce that $\zeta(w_0)$ is a generator of $\Lambda^{mn}V^m$. In particular $\zeta(w_0) \neq 0$.

Let $\tau_L\colon \mathrm{End}(V) \to \mathrm{End}(V^m)$ denote the direct sum of m copies of the tautological representation of $\mathrm{End}(V)$ on V. Then

$$\delta(\sigma_L(g)f) = \tau_L(g)\delta(f)$$

for all $f, g \in \mathrm{End}(V)$. Taking the mn-th exterior power, we get

$$\zeta\big((\Lambda^{mn}\sigma_L)(g)w\big) = \det(\tau_L(g))\zeta(w) = \det(g)^m\zeta(w)$$

for all $g \in \mathrm{End}(V)$ and $w \in W$. □

Lemma 6.4. — $\rho_L(\mathbf{G}(\mathbb{C}))w_0 = \{w \in (\Lambda^{mn}\sigma_L)(Q)w_0 : \zeta(w) = \zeta(w_0)\}$.

Proof. — If $g \in \mathbf{G}(\mathbb{C})$, then we can write $g = sg'$ where $s \in \mathbb{C}^\times$ and $g' \in \mathbf{Sp}_n(\mathbb{C})$. (Choose s to be a square root of $\nu(g)$.) Then $g' \in Q$, $\rho_L(g) = (\Lambda^{mn}\sigma_L)(g')$ and

$$\zeta(\rho_L(g)w_0) = \det((\Lambda^{mn}\sigma_L)(g'))^m\zeta(w_0) = \zeta(w_0),$$

so $\rho_L(g)w_0$ is in $\{w \in (\Lambda^{mn}\sigma_L)(Q)w_0 : \zeta(w) = \zeta(w_0)\}$.

Conversely, if $w = (\Lambda^{mn}\sigma_L)(g)w_0$ for some $g \in Q$ and $\zeta(w) = \zeta(w_0)$, then

$$\det(g)^m\zeta(w_0) = \zeta\big((\Lambda^{mn}\sigma_L)(g)w_0\big) = \zeta(w) = \zeta(w_0).$$

Since $\zeta(w_0) \neq 0$, we deduce that $\det(g)^m = 1$. In particular $\det(g) \neq 0$. Together with $g \in Q$, this implies that $g \in \mathbf{GSp}_n(\mathbb{C})$. Furthermore,

$$\rho_L(g) = (\Lambda^{mn}\sigma_L)(g) \otimes \det(g)^m = (\Lambda^{mn}\sigma_L)(g),$$

so $\rho_L(g)w_0 = w$. Thus $w \in \rho_L(\mathbf{G}(\mathbb{C}))w_0$. □

Thus $\rho_L(\mathbf{G}(\mathbb{C}))w_0$ is Zariski closed in $W_{\mathbb{C}}$, so by [BHC62, Prop. 6.3] $\rho_L(\mathbf{G}(\mathbb{R}))w_0$ is closed in $W_{\mathbb{R}}$ in the real topology.

7. Arithmetic bound for the representation

In this section, we bound the lengths of the vectors v_u of [DO22, Th. 1.2] (here renamed w_u), when applied to the representation ρ_L defined in section 6. This bound is arithmetic in nature, being in terms of discriminants of orders in $\mathbb{Q}$-division algebras. The argument generalises [DO22, §5.5] and Theorem 1.2 plays the role of [DO22, Lem. 5.7].

Proposition 7.1. — *Let d, e and m be positive integers such that dm is even. Let $n = d^2em$. Let $L = \mathbb{Z}^n$ and let $\phi\colon L \times L \to \mathbb{Z}$ be the standard symplectic form as in section* 5.A. *Let $\mathbf{G} = \mathbf{GSp}(L_{\mathbb{Q}}, \phi) = \mathbf{GSp}_{n,\mathbb{Q}}$ and let $\Gamma = \mathbf{Sp}_n(\mathbb{Z})$. Let $\mathbf{H}_0$ be the subgroup of $\mathbf{G}$ defined in* (5.2). *Let W, $\Lambda \subset W$, $\rho_L, \rho_R\colon \mathbf{G} \to \mathbf{GL}(W)$ and $w_0 \in \Lambda$ be as in Proposition* 6.1.

Then there exist positive constants C_{19}, C_{20}, C_{21} and C_{22} such that, for each $u \in \mathbf{G}(\mathbb{R})$, if the group $\mathbf{H}_u = u\mathbf{H}_{0,\mathbb{R}}u^{-1}$ is defined over $\mathbb{Q}$ and $L_{\mathbb{Q}}$ is irreducible as a representation of $\mathbf{H}_u$ over $\mathbb{Q}$, then

(a) *there exists $w_u \in \mathrm{Aut}_{\rho_L(\mathbf{G})}(\Lambda_{\mathbb{R}})w_0$ such that $\rho_L(u)w_u \in \Lambda$ and*

$$|w_u| \leqslant C_{19}|\mathrm{disc}(R_u)|^{C_{20}};$$

(b) *there exists $\gamma \in \Gamma$ and $h \in \mathbf{H}_0(\mathbb{R})$ such that*

$$\|\gamma uh\| \leqslant C_{21}|\mathrm{disc}(R_u)|^{C_{22}},$$

where R_u denotes the ring $\mathrm{End}_{\mathbf{H}_u}(L) \subset \mathrm{M}_n(\mathbb{Z})$.

Note that $L_{\mathbb{Q}}$ is irreducible as a representation of $\mathbf{H}_u$ if and only if $R_{u,\mathbb{Q}}$ is a division algebra. Because $R_{u,\mathbb{R}}$ is $\mathbf{G}(\mathbb{R})$-conjugate to $\mathrm{End}_{\mathbf{H}_0}(L_{\mathbb{R}})$, $R_{u,\mathbb{Q}}$ is an $\mathbb{R}$-split algebra with positive involution. Hence whenever $R_{u,\mathbb{Q}}$ is a division algebra, it must be of type I or II in the Albert classification, and d must equal 1 or 2 for Proposition 7.1 to be non-vacuous.

Let $V = L_{\mathbb{Q}} = \mathbb{Q}^{2g}$. Define $D_0 = \mathrm{M}_d(\mathbb{Q})^e$, $\iota_0\colon D_0 \to \mathrm{M}_n(\mathbb{Q})$, $t\colon D_0 \to D_0$ and $\psi_0\colon V \times V \to D_0$ as in Section 5.B.

By Lemma 3.5, we can choose a D_0-basis $w_1, \ldots, w_m$ for V which is either symplectic or unitary depending on the type of D_0.

Define a symmetric $\mathbb{Q}$-bilinear form $\sigma_0 \colon V \times V \to \mathbb{Q}$ by

$$\sigma_0(x_1 w_1 + \cdots + x_m w_m, y_1 w_1 + \cdots + y_m w_m)) = \mathrm{Trd}_{D_0/\mathbb{Q}}\Big(\sum_{i=1}^{m} x_i y_i^t\Big)$$

for all $x_1, \ldots, x_m, y_1, \ldots, y_m \in D_0$. This bilinear form is positive definite because t is a positive involution. In fact, a lengthy calculation shows that σ_0 is the standard Euclidean inner product on $V = \mathbb{Q}^n$, but we shall not need this fact.

As in the statement of Proposition 7.1, let $u \in \mathbf{G}(\mathbb{R})$ be such that $\mathbf{H}_u = u\mathbf{H}_{0,\mathbb{R}}u^{-1}$ is defined over $\mathbb{Q}$ and V is irreducible as a representation of $\mathbf{H}_u$. Let $D = \mathrm{End}_{\mathbf{H}_u}(V)$, which is a division algebra of type I or II depending on whether $d = 1$ or 2. By construction, V is a left D-vector space of dimension m.

Because $\iota_0(D_0) = \mathrm{End}_{\mathbf{H}_0}(V)$ and $\mathbf{H}_u = u\mathbf{H}_{0,\mathbb{R}}u^{-1}$, we have

$$D = u\iota_0(D_{0,\mathbb{R}})u^{-1} \cap \mathrm{M}_n(\mathbb{Q}).$$

Let $\alpha \colon D_{0,\mathbb{R}} \to D_{\mathbb{R}}$ be the isomorphism of $\mathbb{R}$-algebras

$$\alpha(d) = u\iota_0(d)u^{-1}.$$

Let $\dagger = \alpha \circ t \circ \alpha^{-1}$, which is a positive involution of $D_{\mathbb{R}}$. A calculation using the fact that $u \in \mathbf{G}(\mathbb{R}) = \mathbf{GSp}_n(\mathbb{R})$ shows that ϕ is $(D_{\mathbb{R}}, \dagger)$-compatible, that is, $\dagger$ is the adjoint involution of $D_{\mathbb{R}}$ with respect to ϕ. This has two consequences:

(1) $\dagger$ is defined over $\mathbb{Q}$, that is, $\dagger$ is an involution of D and not just of $D_{\mathbb{R}}$.

(2) There is a non-degenerate $(D, \dagger)$-skew-Hermitian form $\psi \colon V \times V \to D$ such that $\phi = \mathrm{Trd}_{D/\mathbb{Q}} \psi$, thanks to Lemma 3.1.

We are thus in a position to apply Theorem 1.2 (with $R = R_u = \mathrm{Stab}_D(L)$). Let $v_1, \ldots, v_m$ be the resulting weakly symplectic or weakly unitary D-basis for V.

Define a $\mathbb{Q}$-bilinear form $\sigma \colon V \times V \to \mathbb{Q}$ by

$$\sigma\Big(\sum_{i=1}^{m} x_i v_i, \sum_{i=1}^{m} y_i v_i\Big) = \mathrm{Trd}_{D/\mathbb{Q}}\Big(\sum_{i=1}^{m} x_i y_i^{\dagger}\Big)$$

for all $x_1, \ldots, x_m, y_1, \ldots, y_m \in D$.

Lemma 7.2. — *The bilinear form σ is symmetric and positive definite. It takes integer values on $R_u v_1 + \cdots + R_u v_m$ and it satisfies*

$$|\mathrm{disc}(R_u v_1 + \cdots + R_u v_m, \sigma)| = d^{-d^2 em} |\mathrm{disc}(R_u)|^m.$$

Proof. — The form σ is symmetric because $\mathrm{Trd}_{D/\mathbb{Q}}(xy^{\dagger}) = \mathrm{Trd}_{D/\mathbb{Q}}(yx^{\dagger})$ and it is positive definite because $\dagger$ is a positive involution of D.

For each $a \in R_u$ and $y \in L$, the map

$$x \longmapsto \phi(x, a^{\dagger}y) = \phi(ax, y)$$

is $\mathbb{Z}$-linear and maps L into $\mathbb{Z}$. Since ϕ is a perfect pairing on L, this implies that $a^{\dagger}y \in L$ for all $y \in L$. Hence $a^{\dagger} \in \mathrm{Stab}_D(L) = R_u$.

Thus if $x_1, \ldots, x_m, y_1, \ldots, y_m \in R_u$, then each $x_i y_i^{\dagger}$ is in R_u and so $\mathrm{Trd}_{D/\mathbb{Q}}(x_i y_i^{\dagger}) \in \mathbb{Z}$. Hence $\sigma(\sum x_i v_i, \sum y_i v_i) \in \mathbb{Z}$.

For each i, the restriction of σ to $R_u v_i$ is isometric to the inner product associated with $|\cdot|_D$ on R_u. Hence $|\mathrm{disc}(R_u v_i, \sigma)| = d^{-d^2 e}|\mathrm{disc}(R_u)|$ and so

$$|\mathrm{disc}(R_u v_1 + \cdots + R_u v_m, \sigma)| = d^{-d^2 em}|\mathrm{disc}(R_u)|^m. \qquad \square$$

Lemma 7.3. — *There exists an $\mathbb{R}$-linear map $\theta\colon V_{\mathbb{R}} \to V_{\mathbb{R}}$ with the following properties:*

(i) *$\theta(\alpha(a)x) = \iota_0(a)\theta(x)$ for all $a \in D_{0,\mathbb{R}}$, $x \in V_{\mathbb{R}}$;*

(ii) *$\psi = \alpha \circ \theta^* \psi_0$;*

(iii) *$\sigma_0(\theta(x), \theta(x)) \leqslant C_{23}|\mathrm{disc}(R_u)|^{C_{24}}\sigma(x,x)$ for all $x \in V_{\mathbb{R}}$, where the constants depend only on d, e and m (and not on u).*

Proof. — Use Lemma 3.4 to choose $s_1, \ldots, s_m \in D_{\mathbb{R}}^\times$ such that $s_1^{-1}v_1, \ldots, s_m^{-1}v_m$ is a symplectic or α-unitary $D_{\mathbb{R}}$-basis for $V_{\mathbb{R}}$.

Define $\theta\colon V_{\mathbb{R}} \to V_{\mathbb{R}}$ by

$$\theta(x_1 v_1 + \cdots + x_m v_m) = \iota_0(\alpha^{-1}(x_1 s_1))w_1 + \cdots + \iota_0(\alpha^{-1}(x_m s_m))w_m$$

for all $x_1, \ldots, x_m \in D_{\mathbb{R}}$.

– Claim (i) holds because $\alpha\colon D_{0,\mathbb{R}} \to D_{\mathbb{R}}$ is a ring homomorphism.

– Claim (ii) holds because $s_1^{-1}v_1, \ldots, s_m^{-1}v_m$ is a symplectic or α-unitary $D_{\mathbb{R}}$-basis for $V_{\mathbb{R}}$ while $w_1, \ldots, w_m$ is a symplectic or unitary $D_{0,\mathbb{R}}$-basis for $D_{0,\mathbb{R}}^m$. Thus

$$\alpha(\psi_0(\theta(s_i^{-1}v_i), \theta(s_j^{-1}v_j))) = \alpha(\psi_0(w_i, w_j)) = \psi(s_i^{-1}v_i, s_j^{-1}v_j) \quad \text{for all } i, j.$$

– For claim (iii): for every $x = x_1 v_1 + \cdots + x_m v_m \in V_{\mathbb{R}}$, where $x_1, \ldots, x_m \in D_{\mathbb{R}}$,

$$\begin{aligned}
\sigma_0(\theta(x), \theta(x)) &= \mathrm{Trd}_{D_{0,\mathbb{R}}/\mathbb{R}}\Big(\sum_{i=1}^m \alpha^{-1}(x_i s_i)\alpha^{-1}(x_i s_i)^t\Big) \\
&= \sum_{i=1}^m \mathrm{Trd}_{D_{0,\mathbb{R}}/\mathbb{R}}\big(\alpha^{-1}(x_i s_i s_i^\dagger x_i^\dagger)\big) \\
&= \sum_{i=1}^m \mathrm{Trd}_{D_{\mathbb{R}}/\mathbb{R}}(x_i s_i s_i^\dagger x_i^\dagger) = \sum_{i=1}^m |x_i s_i|_D^2 \\
&\leqslant \sum_{i=1}^m |x_i|_D^2 |s_i|_D^2 \leqslant \Big(\max_{i=1,\ldots,m} |s_i|_D^2\Big)\sigma(x,x).
\end{aligned} \tag{7.1}$$

Thanks to Lemma 3.4 and Theorem 1.2(iv), we have

$$\max_{i=1,\ldots,m} |s_i|_D^2 \leqslant (de)^{1/2} \max_{i,j=1,\ldots,m} |\psi(v_i, v_j)|_D \leqslant C_{25}|\mathrm{disc}(R_u)|^{C_{26}},$$

where the constants depend only on d, e and m. Combined with (7.1), this proves claim (iii). $\square$

Lemma 7.4. — *Let $h = u^{-1}\theta^{-1}\colon V_{\mathbb{R}} \to V_{\mathbb{R}}$. Then $uh = \theta^{-1} \in \mathbf{Sp}_n(\mathbb{R})$ and $h \in \mathbf{H}_0(\mathbb{R})$.*

Proof. — Firstly, $\theta \in \mathbf{Sp}_n(\mathbb{R})$ by the following calculation, which relies on Lemma 7.3(ii):

$$\begin{aligned}\theta^*\phi = \theta^*(\mathrm{Trd}_{D_{0,\mathbb{R}}/\mathbb{R}} \circ \psi_0) &= \theta^*(\mathrm{Trd}_{D_{\mathbb{R}}/\mathbb{R}} \circ \alpha \circ \psi_0) \\ &= \mathrm{Trd}_{D_{\mathbb{R}}/\mathbb{R}} \circ \alpha \circ (\theta^*\psi_0) = \mathrm{Trd}_{D_{\mathbb{R}}/\mathbb{R}} \circ \psi = \phi.\end{aligned}$$

Since $\mathbf{Sp}_n(\mathbb{R}) \subset \mathbf{GSp}_n(\mathbb{R}) = \mathbf{G}(\mathbb{R})$ and $u \in \mathbf{G}(\mathbb{R})$, it follows that $h \in \mathbf{G}(\mathbb{R})$. By definition, $\mathbf{H}_0 = Z_{\mathbf{G}}(\iota_0(D_0))$ and so it remains to prove that h commutes with the action of D_0 on V. For $a \in D_{0,\mathbb{R}}$ and $x \in V_{\mathbb{R}}$, we have

$$\begin{aligned}h(\iota_0(a)x) = u^{-1}\theta^{-1}(\iota_0(a)x) &= u^{-1}\alpha(a)\theta^{-1}(x) \\ &= \iota_0(a)u^{-1}\theta^{-1}(x) = \iota_0(a)h(x)\end{aligned}$$

where we use Lemma 7.3(i) and the fact that $\alpha(a) = u\iota_0(a)u^{-1}$ (from the definition of α). Thus h commutes with all $a \in \iota_0(D_0)$. □

Lemma 7.5. — *There exists a $\mathbb{Z}$-basis $\{e'_1, \dots, e'_n\}$ for L such that the coordinates of the vectors $\theta(e'_1), \dots, \theta(e'_n)$ in $V_{\mathbb{R}} = \mathbb{R}^n$ are polynomially bounded in terms of $|\mathrm{disc}(R_u)|$.*

Proof. — Let $\lambda_1, \dots, \lambda_n$ denote the successive minima of $R_u v_1 + \cdots + R_u v_m$ with respect to σ. By Theorem 2.4 and Lemma 7.2, we have

$$\lambda_1\lambda_2\cdots\lambda_n \leqslant \gamma_{d^2em}^{n/2}\,\mathrm{covol}(R_u v_1 + \cdots + R_u v_m) \leqslant C_{27}|\mathrm{disc}(R_u)|^{-m},$$

where C_{27} depends only on d, e and m.

For each i, $\lambda_i^2 = \sigma(v, v)$ for some $v \in R_u v_1 + \cdots + R_u v_m$ and so $\lambda_i \geqslant 1$ by Lemma 7.2. We deduce that, for each i,

$$\lambda_i \leqslant C_{27}|\mathrm{disc}(R_u)|^{-m}.$$

Let $\lambda'_1, \dots, \lambda'_n$ denote the successive minima of L with respect to σ. Since $R_u v_1 + \cdots + R_u v_m \subset L$, $\lambda'_i \leqslant \lambda_i$ for each i. By [Wey40, Th. 4], there exists a $\mathbb{Z}$-basis $e'_1, \dots, e'_n$ for L such that

$$\sqrt{\sigma(e'_i, e'_i)} \leqslant C_{28}\lambda'_i,$$

where C_{28} depends only on n. Combining the above inequalities, we obtain

$$\sigma(e'_i, e'_i) \leqslant C_{32}|\mathrm{disc}(R_u)|^{-2m}.$$

Combining this with Lemma 7.3(iii), we obtain that

$$\sigma_0(\theta(e'_i), \theta(e'_i)) \leqslant C_{23}|\mathrm{disc}(R_u)|^{C_{24}}\sigma(e'_i, e'_i) \leqslant C_{30}|\mathrm{disc}(R_u)|^{C_{31}}$$

for some constants C_{30}, C_{31} independent of $u \in \mathbf{G}(\mathbb{R})$. Since σ_0 is a fixed positive definite quadratic form on $V_{\mathbb{R}}$, this implies that the coordinates of the vectors $\theta(e'_1), \dots, \theta(e'_n)$ are likewise bounded by a polynomial in $|\mathrm{disc}(R_u)|$. □

Let γ' be the matrix in $\mathbf{GL}_n(\mathbb{Z})$ which maps the vectors $e'_1, \dots, e'_n$ to the standard basis of $L = \mathbb{Z}^n$.

Lemma 7.6. — *The entries of the matrices $\gamma' uh, (\gamma' uh)^{-1} \in \mathbf{GL}_n(\mathbb{R})$ are polynomially bounded in terms of* $\mathrm{disc}(R_u)$.

Proof. — Let $A = \gamma' uh = \gamma'\theta^{-1} \in \mathbf{GL}_n(\mathbb{R})$. Observe that A maps $\theta(e_1'), \dots, \theta(e_n')$ to the standard basis. In other words, the entries of A^{-1} are the coordinates of $\theta(e_1'), \dots, \theta(e_n')$ and so are bounded by Lemma 7.5.

By Lemma 7.4, $\det(uh) = 1$, while $|\det(\gamma')| = 1$ since $\gamma' \in \mathbf{GL}_n(\mathbb{Z})$. Hence $|\det(A)| = 1$. By Cramer's rule, each entry of A is a fixed polynomial in the entries of A^{-1}, multiplied by $\det(A)$. We conclude that the entries of A are polynomially bounded in terms of $\mathrm{disc}(R_u)$. □

We now show that we can modify $\gamma' \in \mathbf{GL}_n(\mathbb{Z})$ to obtain $\gamma \in \mathbf{Sp}_n(\mathbb{Z})$, with a similar bound on γuh. This establishes Proposition 7.1(b), and we will subsequently use it to prove Proposition 7.1(a).

Lemma 7.7. — *There exists $\gamma \in \Gamma = \mathbf{Sp}_n(\mathbb{Z})$ such that the entries of γuh and $(\gamma uh)^{-1}$ are polynomially bounded in terms of $|\mathrm{disc}(R_u)|$.*

Proof. — Let $e_1, \dots, e_n$ denote the standard basis of $L = \mathbb{Z}^n$.

According to Lemma 7.4, $uh \in \mathbf{Sp}_n(\mathbb{R})$. Consequently,

$$\phi(\gamma'^{-1}e_i, \gamma'^{-1}e_j) = \phi((uh)^{-1}\gamma'^{-1}e_i,\ (uh)^{-1}\gamma'^{-1}e_j) \text{ for all } i, j \in \{1, \dots, n\}.$$

By Lemma 7.6, the entries of $(uh)^{-1}\gamma'^{-1}$ are polynomially bounded in terms of $|\mathrm{disc}(R_u)|$, and hence the same is true of the values $\phi(\gamma'^{-1}e_i, \gamma'^{-1}e_j)$.

Hence, by [Orr15, Lem. 4.3], there exists a symplectic $\mathbb{Z}$-basis $\{f_1, \dots, f_n\}$ for (L, ϕ) whose coordinates with respect to $\{\gamma'^{-1}e_1, \dots, \gamma'^{-1}e_n\}$ are polynomially bounded in terms of $|\mathrm{disc}(R_u)|$. Applying γ', we deduce that the coordinates of $\gamma' f_1, \dots, \gamma' f_n$ with respect to the standard basis are polynomially bounded.

Let $\gamma \in \mathbf{GL}_n(\mathbb{Z})$ be the matrix such that $e_i = \gamma f_i$ for each $i = 1, \dots, n$. Since $\{f_1, \dots, f_n\}$ is a symplectic basis, we have $\gamma \in \Gamma$. We have just shown that the coordinates of $\gamma' f_i = \gamma'\gamma^{-1}e_i$ are polynomially bounded for each i. In other words, the entries of the matrix $\gamma'\gamma^{-1}$ are polynomially bounded in terms of $|\mathrm{disc}(R_u)|$.

Multiplying $(\gamma' uh)^{-1}$ by $\gamma'\gamma^{-1}$ and applying Lemma 7.6, we deduce that the entries of $(\gamma uh)^{-1}$ are polynomially bounded in terms of $|\mathrm{disc}(R_u)|$. Thanks to Lemma 7.4, $|\det(\gamma uh)| = 1$, so it follows that the entries of γuh are also polynomially bounded in terms of $|\mathrm{disc}(R_u)|$. □

Let $S_u = \mathrm{End}_{R_u}(L) = uE_{0,\mathbb{R}}u^{-1} \cap \mathrm{M}_n(\mathbb{Z})$, where E_0 is defined in (6.1). By Proposition 6.1(v), there exists $d_u \in \mathbb{R}_{>0}$ such that

$$d_u \rho_R(u)\rho_L(u)w_0 \in \Lambda \quad \text{and} \quad d_u \leqslant C_{18}|\mathrm{disc}(S_u)|^{1/2}.$$

In order to prove Proposition 7.1(a), we shall use the vector

$$w_u = d_u \rho_R(\gamma u) w_0 \in W_{\mathbb{R}}.$$

Observe first that $d_u\rho_R(\gamma u) \in \mathrm{Aut}_{\rho_L(\mathbf{G})}(\Lambda_{\mathbb{R}})$ thanks to Proposition 6.1(iii), and that $\rho_L(u)w_u = \rho_R(\gamma)d_u\rho_R(u)\rho_L(u)w_0$ is in Λ thanks to Proposition 6.1(iv) . Hence w_u satisfies the qualitative conditions of Proposition 7.1(a), and it only remains to prove the bound for $|w_u|$.

Lemma 7.8. — $|w_u| \leqslant C_{32}|\mathrm{disc}(R_u)|^{C_{33}}$.

Proof. — According to Proposition 6.1(i), $\mathbf{H}_{d,e,m} = \mathrm{Stab}_{\rho_R(\mathbf{G})}(w_0)$. Therefore

$$w_u = d_u\rho_R(\gamma u)w_0 = d_u\rho_R(\gamma uh)w_0.$$

The homomorphism $\rho_R\colon \mathbf{G} \to \mathbf{GL}(W)$ is given by fixed polynomials in the entries and inverse determinant. Since the entries of γuh and $\det(\gamma uh)^{-1}$ are bounded by Lemma 7.7, we deduce that the entries of $\rho_R(\gamma uh)$ are likewise polynomially bounded in terms of $\mathrm{disc}(R_u)$.

Meanwhile, by definition, d_u is polynomially bounded in terms of $\mathrm{disc}(S_u)$. By Lemma 2.10, $\mathrm{disc}(S_u)$ is polynomially bounded in terms of $\mathrm{disc}(R_u)$. We conclude that $|w_u|$ is polynomially bounded in terms of $|\mathrm{disc}(R_u)|$, as required. □

8. Cases of Zilber–Pink

In this section, we prove Theorems 1.3 and 1.5. The proofs follow closely [DO22, §6]. We refer to notation and terminology from [Orr18, §2.2 and 2.4].

8.A. Proof of Theorem 1.3.

— In fact, instead of proving Theorem 1.3, we will prove the following, more general theorem. (Recall that, by Proposition 5.5, for $g \geqslant 3$, all proper special subvarieties of PEL type of $\mathcal{A}_g$ have codimension at least 2.)

Theorem 8.1. — *Let $g \geqslant 3$ and let C be an irreducible algebraic curve in $\mathcal{A}_g$. Let S denote the smallest special subvariety of $\mathcal{A}_g$ containing C. Let Ω denote the set of special subvarieties of $\mathcal{A}_g$ of simple PEL type I or II of dimension at most $\dim(S)-2$. Let Σ denote the set of points in $\mathcal{A}_g(\mathbb{C})$ which are endomorphism generic in some $Z \in \Omega$.*

If C satisfies Conjecture 8.2, then $C \cap \Sigma$ is finite.

Conjecture 8.2 is the natural generalisation of Conjecture 1.4.

Conjecture 8.2. — *Let C and Σ be as in Theorem 8.1 and let L be a finitely generated subfield of $\mathbb{C}$ over which C is defined. Then there exist positive constants C_{34} and C_{35} such that*

$$\#\,\mathrm{Aut}(\mathbb{C}/L) \cdot s \geqslant C_{34}|\,\mathrm{disc}(\mathrm{End}(A_s))|^{C_{35}}$$

for all $s \in C \cap \Sigma$.

Let $L = \mathbb{Z}^{2g}$ and let $\phi : L \times L \to \mathbb{Z}$ be the standard symplectic form as in section 5.A. Let $\mathbf{G} = \mathbf{GSp}(L_{\mathbb{Q}}, \psi) = \mathbf{GSp}_{2g}$ and let $\Gamma = \mathbf{Sp}_{2g}(\mathbb{Z})$. Define $h_0\colon \mathbb{S} \to \mathbf{G}_{\mathbb{R}}$ as in (5.1) and let X^+ denote the $\mathbf{G}(\mathbb{R})$-conjugacy class of h_0 in $\mathrm{Hom}(\mathbb{S}, \mathbf{G}_{\mathbb{R}})$. Then $(\mathbf{G}, X^+)$ is a Shimura datum component and so $\mathrm{Stab}_{\mathbf{G}(\mathbb{R})}(h_0) = \mathbb{R}^\times K_\infty^+$ where K_∞^+ is a maximal compact subgroup of $\mathbf{G}(\mathbb{R})^+$ [Mil05, Chap. 6].

Let $(\mathbf{P}, \mathbf{S}, K_\infty)$ be a Siegel triple for $\mathbf{G}$, as defined in [Orr18, §2B], where K_∞ is a maximal compact subgroup of $\mathbf{G}(\mathbb{R})$ such that $K_\infty^+ = \mathbf{G}(\mathbb{R})^+ \cap K_\infty$. By the results of Borel quoted in [Orr18, §2D], there exists a Siegel set $\mathfrak{S} \subset \mathbf{G}(\mathbb{R})$ with respect to

$(\mathbf{P}, \mathbf{S}, K_\infty)$ and a finite set $C_{\mathbf{G}} \subset \mathbf{G}(\mathbb{Q})$ such that $\mathcal{F}_{\mathbf{G}} = C_{\mathbf{G}}\mathfrak{S}$ is a fundamental set for Γ in $\mathbf{G}(\mathbb{R})$.

Let $\mathcal{F} = (\mathcal{F}_{\mathbf{G}} \cap \mathbf{G}(\mathbb{R})^+)h_0$. Since $\Gamma \subset \mathbf{G}(\mathbb{R})^+$, $\mathcal{F}$ is a fundamental set in X^+ for Γ. If we denote by $\pi : X^+ \to \mathcal{A}_g$ the uniformising map, then $\pi|_{\mathcal{F}}$ is definable in the o-minimal structure $\mathbb{R}_{\mathrm{an,exp}}$ (see [PS10] for the original result and [KUY16] for a formulation in notations more similar to ours).

As explained in section 1.D, Σ is the union of sets $\Sigma_{d,e,m}$, where d, e, m are positive integers satisfying $d^2em = 2g$, $d = 1$ or 2 and dm is even. Since there are only finitely many choices for such d, e, m (given g), in order to prove Theorem 8.1, it suffices to prove that $C \cap \Sigma_{d,e,m}$ is finite for each d, e, m.

From now on, we fix such integers d, e and m. Let $\mathbf{H}_0 \subset \mathbf{G}$ be the group defined in (5.2) associated with these parameters. Let $X_0^+ = \mathbf{H}_0(\mathbb{R})^+h_0$, so that $(\mathbf{H}_0, X_0^+)$ is the unique Shimura subdatum component of $(\mathbf{G}, X^+)$ given by Lemma 5.2.

By Propositions 6.1 and 7.1, there exists a finitely generated, free $\mathbb{Z}$–module Λ, a representation $\rho_L : \mathbf{G} \to \mathbf{GL}(\Lambda_{\mathbb{Q}})$ such that Λ is stabilised by $\rho_L(\Gamma)$, a vector $w_0 \in \Lambda$ and positive constants C_{19} and C_{20} such that:

(i) $\mathrm{Stab}_{\mathbf{G},\rho_L}(w_0) = \mathbf{H}_0$;

(ii) the orbit $\rho_L(\mathbf{G}(\mathbb{R}))w_0$ is closed in $\Lambda_{\mathbb{R}}$;

(iii) for each $u \in \mathbf{G}(\mathbb{R})$, if the group $\mathbf{H}_u = u\mathbf{H}_{0,\mathbb{R}}u^{-1}$ is defined over $\mathbb{Q}$ and $L_{\mathbb{Q}}$ is irreducible as a representation of $\mathbf{H}_u$ over $\mathbb{Q}$, then there exists $w_u \in \mathrm{Aut}_{\rho_L(\mathbf{G})}(\Lambda_{\mathbb{R}})w_0$ such that $\rho_L(u)w_u \in \Lambda$ and

$$|w_u| \leqslant C_{19}|\mathrm{disc}(R_u)|^{C_{20}},$$

where R_u denotes the ring $\mathrm{End}_{\mathbf{H}_u}(L) \subset \mathrm{M}_{2g}(\mathbb{Z})$.

By [DO22, Th. 1.2], there exist positive constants C_{36} and C_{37} with the following property: for every $u \in \mathbf{G}(\mathbb{R})$ and $w_u \in \mathrm{Aut}_{\rho_L(\mathbf{G})}(\Lambda_{\mathbb{R}})w_0$ such that $\mathbf{H}_u = u\mathbf{H}_{0,\mathbb{R}}u^{-1}$ is defined over $\mathbb{Q}$ and $\rho_L(u)w_u \in \Lambda$, there exists a fundamental set for $\Gamma \cap \mathbf{H}_u(\mathbb{R})$ in $\mathbf{H}_u(\mathbb{R})$ of the form

$$B_u\mathcal{F}_{\mathbf{G}}u^{-1} \cap \mathbf{H}_u(\mathbb{R}),$$

where $B_u \subset \Gamma$ is a finite set such that

$$|\rho_L(b^{-1}u)w_u| \leqslant C_{36}|w_u|^{C_{37}}$$

for every $b \in B_u$.

For any $w \in \Lambda_{\mathbb{R}}$, we write $\mathbf{G}(w)$ for the real algebraic group $\mathrm{Stab}_{\mathbf{G}_{\mathbb{R}},\rho_L}(w)$. Fixing a basis for Λ, we may refer to the height $\mathrm{H}(w)$ of any $w \in \Lambda$ (namely, the maximum of the absolute values of its coordinates with respect to this basis.)

Lemma 8.3. — *Let $P \in \Sigma_{d,e,m}$. There exists $z \in \pi^{-1}(P) \cap \mathcal{F}$ and*

$$w \in \mathrm{Aut}_{\rho_L(\mathbf{G})}(\Lambda_{\mathbb{R}})\rho_L(\mathbf{G}(\mathbb{R})^+)w_0 \cap \Lambda$$

such that $z(\mathbb{S}) \subset \mathbf{G}(w)$ and

$$\mathrm{H}(w) \leqslant C_{36}C_{19}^{C_{37}}|\mathrm{disc}(R)|^{C_{20}C_{37}},$$

where $R = \mathrm{End}(A_P) \cong \mathrm{End}_{\mathbf{G}(w)}(L) \subset \mathrm{M}_{2g}(\mathbb{Z})$.

Proof. — Let $z' \in \pi^{-1}(P) \cap \mathcal{F}$. Since $P \in \Sigma_{d,e,m}$, it is an endomorphism generic point of a special subvariety $S \subset \mathcal{A}_g$ of simple PEL type I or II with parameters d, e, m. Therefore, there is a Shimura subdatum component $(\mathbf{H}, Y^+) \subset (\mathbf{G}, X^+)$ of simple PEL type I or II such that $\pi(Y^+) = S$ and $z' \in Y^+$. (In particular, $z'(\mathbb{S}) \subset \mathbf{H}_{\mathbb{R}}$.) By Lemma 5.1, $\mathbf{H}_{\mathbb{R}} = u\mathbf{H}_{0,\mathbb{R}}u^{-1}$ for some $u \in \mathbf{G}(\mathbb{R})^+$, and so we write $\mathbf{H}_u = \mathbf{H}$. By Corollary 5.3,

$$Y^+ = uX_0^+ = u\mathbf{H}_0(\mathbb{R})^+ h_0 = \mathbf{H}_u(\mathbb{R})^+ uh_0.$$

Let $R_u = \mathrm{End}_{\mathbf{H}_u}(L)$. Since $\mathbf{H}_u$ is the general Lefschetz group of S, R_u is the generic endomorphism ring of S and, hence, isomorphic to $\mathrm{End}(A_P)$. Since S is a special subvariety of simple PEL type, $R_{u,\mathbb{Q}}$ is a division algebra. Hence $L_{\mathbb{Q}}$ is irreducible as a representation of $\mathbf{H}_u$.

By Proposition 7.1(a), there exists $w_u \in \mathrm{Aut}_{\rho_L(\mathbf{G})}(\Lambda_{\mathbb{R}})w_0$ such that

$$\rho_L(u)w_u \in \Lambda \quad \text{and} \quad |w_u| \leqslant C_{19}|\mathrm{disc}(R_u)|^{C_{20}}.$$

Hence, by [DO22, Th. 1.2], there exists a fundamental set for $\Gamma \cap \mathbf{H}_u(\mathbb{R})$ in $\mathbf{H}_u(\mathbb{R})$ of the form

$$\mathcal{F}_u = B_u \mathcal{F}_{\mathbf{G}} u^{-1} \cap \mathbf{H}_u(\mathbb{R}),$$

where $B_u \subset \Gamma$ is a finite set such that

$$|\rho_L(b^{-1}u)w_u| \leqslant C_{36}|w_u|^{C_{37}}$$

for every $b \in B_u$. Therefore, we can write $z' \in \mathbf{H}_u(\mathbb{R})^+ uh_0$ as

$$z' = \gamma b f u^{-1} \cdot uh_0$$

for some $\gamma \in \Gamma \cap \mathbf{H}_u(\mathbb{R})$, $b \in B_u$, and $f \in \mathcal{F}_{\mathbf{G}}$.

Set

$$z = b^{-1}\gamma^{-1}z' = fh_0 \in \mathcal{F}_{\mathbf{G}} h_0 \cap X^+ = \mathcal{F},$$

where the last equality uses the fact that $\mathrm{Stab}_{\mathbf{G}(\mathbb{R})}(h_0) \subset \mathbf{G}(\mathbb{R})^+$. Since $b, \gamma \in \Gamma$, we obtain $z \in \pi^{-1}(P) \cap \mathcal{F}$.

Let $w = \rho_L(b^{-1}u)w_u$. As in [DO22, Prop. 6.3], we can show that $z(\mathbb{S}) \subset \mathbf{G}(w)$ and that $\mathbf{G}(w)$ is a Γ-conjugate of $\mathbf{H}_u$, so $R_u \cong \mathrm{End}_{\mathbf{G}(w)}(L)$. Consequently, z and w satisfy the requirements of the lemma. □

We can now deduce Proposition 1.8.

Corollary 8.4. — *Define $\Sigma \subset \mathcal{A}_g$ as in Theorem 1.3. For each $b \in \mathbb{R}$, the points $s \in \Sigma$ such that $|\mathrm{disc}(\mathrm{End}(A_s))| \leqslant b$ belong to only finitely many proper special subvarieties of simple PEL type I or II.*

Proof. — The proof is essentially the same as [DO22, Cor. 6.4]. □

The proof of Theorem 8.1 now proceeds as in [DO22, §6.5] with some modifications, which we outline below (following the notation from [DO22, §6.5] *mutatis mutandis*).

(1) The argument is carried out inside $X^+ \cong \mathcal{H}_g$ instead of $\mathcal{H}_2$.

(2) If $P \in \Sigma_{d,e,m}$, then P is endomorphism generic in some special subvariety $Z \in \Omega$ (where Ω is defined in Theorem 8.1). Then Z is an irreducible component of $\mathcal{M}_R$, where $R = \mathrm{End}(A_P)$ (see definitions in section 1.C). Since $\mathcal{M}_R$ is $\mathrm{Aut}(\mathbb{C})$-invariant and its (analytic) irreducible components are algebraic subvarieties of $\mathcal{A}_g$, for each $\sigma \in \mathrm{Aut}(\mathbb{C})$, $\sigma(Z)$ is also an irreducible component of $\mathcal{M}_R$. Thus, $\sigma(Z)$ is also a special subvariety of simple PEL type I or II with the same parameters d, e, m. Furthermore, $\dim(\sigma(Z)) = \dim(Z)$, so $\sigma(Z) \in \Omega$. Since $\mathrm{End}(A_{\sigma(P)}) \cong \mathrm{End}(A_P)$, $\sigma(P)$ is endomorphism generic in $\sigma(Z)$, so $\sigma(P) \in \Sigma_{d,e,m}$.

(3) In the definition of the definable set D, we replace $\mathbf{G}(\mathbb{R})$ with $\mathbf{G}(\mathbb{R})^+$. That is, $w \in \mathrm{Aut}_{\rho_L(\mathbf{G})}(\Lambda_{\mathbb{R}})\rho_L(\mathbf{G}(\mathbb{R})^+)w_0$, as in Lemma 8.3. Then

$$g_t \in \mathrm{Aut}_{\rho_L(\mathbf{G})}(\Lambda_{\mathbb{R}})\rho_L(\mathbf{G}(\mathbb{R})^+)$$

for all t. So $g_t^{-1}z_t$ is in the same connected component of X as $z_t \in \mathcal{C} \subset X^+$. We conclude that $g_t^{-1}z_t$ lies on the unique pre-special subvariety of $X^+ \cong \mathcal{H}_g$ associated with $\mathbf{H}_0$, namely, X_0^+ (see Lemma 5.2).

(4) By the inverse Ax–Lindemann conjecture, the smallest algebraic subset of X^+ containing $\widetilde{C}$ is an irreducible component of $\pi^{-1}(S)$, which we call $\widetilde{S}$.

(5) As in the penultimate paragraph of [DO22, §6.5], we choose a complex algebraic subset $\widetilde{B} \subset \mathrm{Aut}_{\rho_L(\mathbf{G})}(\Lambda_{\mathbb{C}})\rho_L(\mathbf{G}(\mathbb{C}))$ of dimension at most 1 whose image under the map $g \mapsto g \cdot w_0$ is B. Here, the map

$$\cdot : \mathrm{Aut}_{\rho_L(\mathbf{G})}(\Lambda_{\mathbb{C}})\rho_L(\mathbf{G}(\mathbb{C})) \times (X^+)^\vee \longrightarrow (X^+)^\vee \cong \mathcal{H}_g^\vee$$

(which is used in [DO22, §6.5], but not explicitly defined there) is given by

$$(a\rho_L(g), x) \longmapsto g \cdot x$$

for each $a \in \mathrm{Aut}_{\rho_L(\mathbf{G})}(\Lambda_{\mathbb{C}})$ and $\rho_L(g) \in \rho_L(\mathbf{G}(\mathbb{C}))$. This is well-defined since

$$\mathrm{Aut}_{\rho_L(\mathbf{G})}(\Lambda_{\mathbb{C}}) \cap \rho_L(\mathbf{G}(\mathbb{C})) \subset \rho_L(Z(\mathbf{G})(\mathbb{C})) \quad \text{and} \quad \ker(\rho_L) \subset Z(\mathbf{G}),$$

and $Z(\mathbf{G})$, the centre of $\mathbf{G}$, acts trivially on $(X^+)^\vee$.

(6) In the final step, we conclude that $\widetilde{B} \cdot (X_0^+)^\vee$ has uncountable intersection with $\widetilde{C}$ and, hence, contains it. Therefore, $\widetilde{S}$ is contained in $\widetilde{B} \cdot (X_0^+)^\vee$, but

$$\dim(\widetilde{B} \cdot (X_0^+)^\vee) \leqslant 1 + \dim(X_0^+) \leqslant \dim(S) - 1,$$

delivering the contradiction.

8.B. Proof of Theorem 1.5. — If C is an algebraic curve over a number field, and $\mathfrak{A} \to C$ is an abelian scheme of even relative dimension g, we say that $s \in C(\overline{\mathbb{Q}})$ is an *exceptional quaternionic point* if $\mathrm{End}(\mathfrak{A}_s) \otimes_{\mathbb{Z}} \mathbb{Q}$ is a non-split totally indefinite quaternion algebra over a totally real field of degree e such that $4e$ does not divide g. Note that these are precisely the points for which:

(i) $\mathfrak{A}_s$ is simple and $D := \mathrm{End}(\mathfrak{A}_s) \otimes \mathbb{Q}$ has type I or II; and

(ii) $\mathfrak{A}_s$ is exceptional in the sense of [DO21, Def. 8.1], that is, D is not isomorphic to a subring of $\mathrm{M}_g(\mathbb{Q})$.

Indeed, if $\mathfrak{A}_s$ is simple, then D is a division algebra and hence embeds into $\mathrm{M}_g(\mathbb{Q})$ if and only if $\dim_{\mathbb{Q}}(D)$ divides g. If D has type I, then $\dim_{\mathbb{Q}}(D)$ always divides g, while if D has type II, then $\dim_{\mathbb{Q}}(D) = 4e$.

In order to prove Theorem 1.5, it suffices to prove the following theorem, by the same argument as in [DO22, §6.7]. This theorem is a direct generalisation of [DO22, Th. 6.5]. Note that the image of $C \to \mathcal{A}_g$ is Hodge generic if and only if the generic Mumford–Tate group of the abelian scheme $\mathfrak{A} \to C$ is $\mathbf{GSp}_{2g,\mathbb{Q}}$.

Theorem 8.5. — *Let C be a irreducible algebraic curve and let $\mathfrak{A} \to C$ be a principally polarised non-isotrivial abelian scheme of even relative dimension g such that the image of the morphism $C \to \mathcal{A}_g$ induced by $\mathfrak{A}$ is Hodge generic.*

Suppose that C and $\mathfrak{A}$ are defined over a number field L and that there exists a smooth curve C', a semi-abelian scheme $\mathfrak{A}' \to C'$ and an open immersion $\iota\colon C \to C'$, all defined over $\overline{\mathbb{Q}}$, such that $\mathfrak{A} \cong \iota^*\mathfrak{A}'$ and, for some point $s_0 \in C'(\overline{\mathbb{Q}}) \smallsetminus C(\overline{\mathbb{Q}})$, the fibre $\mathfrak{A}'_{s_0}$ is a torus.

Then there exist positive constants C_{38} and C_{39} such that, for any exceptional quaternionic point $s \in C$,

$$\#\operatorname{Aut}(\mathbb{C}/L) \cdot s \geqslant C_{38}|\mathrm{disc}(\mathrm{End}(\mathfrak{A}_s))|^{C_{39}}.$$

Proof. — After replacing L by a finite extension, we may assume that C', $\mathfrak{A}' \to C'$ and $\iota\colon C \to C'$ are all defined over L. After replacing C' by its normalisation and $\mathfrak{A}'$ by its pullback to this normalisation, we may assume that C' is smooth. (Note that this step, which is required in order to apply [DO21, Th. 8.2], was erroneously omitted in the proofs of [DO21, Prop. 9.2] and [DO22, Th. 6.5].) Observe that $\mathfrak{A} \to C$ satisfies the conditions of [DO21, Th. 8.2].

Let $s \in C$ be an exceptional quaternionic point. The image of s under the map $C \to \mathcal{A}_g$ induced by $\mathfrak{A} \to C$ is in the intersection between the image of C and a proper special subvariety of PEL type. Since C is a curve defined over $\overline{\mathbb{Q}}$ and special subvarieties of $\mathcal{A}_g$ are defined over $\overline{\mathbb{Q}}$, it follows that $s \in C(\overline{\mathbb{Q}})$.

The remainder of the proof proceeds as in the proof of [DO22, Th. 6.5]. The key ingredients are:

(1) [DO21, Th. 8.2], a height bound for exceptional points of C (including exceptional quaternionic points) which generalises [And89, Chap. X, Th. 1.3];

(2) endomorphism estimates of Masser and Wüstholz [MW94] (a version using present notations is [DO22, Th. 6.6]).

□

References

[AQ97] G. D. Anderson & S.-L. Qiu – "A monotoneity property of the gamma function", *Proc. Amer. Math. Soc.* **125** (1997), no. 11, p. 3355–3362.

[And89] Y. André – *G-functions and geometry*, Aspects of Math., vol. E13, Friedr. Vieweg & Sohn, Braunschweig, 1989.

[BM21] G. Binyamini & D. Masser – "Effective André–Oort for non-compact curves in Hilbert modular varieties", *Comptes Rendus Mathématique* **359** (2021), no. 3, p. 313–321.

[BL04] C. Birkenhake & H. Lange – *Complex abelian varieties*, second ed., Grundlehren Math. Wissen., vol. 302, Springer-Verlag, Berlin, 2004.

[BG06] E. Bombieri & W. Gubler – *Heights in Diophantine geometry*, New Math. Monographs, vol. 4, Cambridge University Press, Cambridge, 2006.

[BHC62] A. Borel & Harish-Chandra – "Arithmetic subgroups of algebraic groups", *Ann. of Math. (2)* **75** (1962), p. 485–535.

[Cas97] J. W. S. Cassels – *An introduction to the geometry of numbers*, Classics in Math., Springer-Verlag, Berlin, 1997.

[CR62] C. W. Curtis & I. Reiner – *Representation theory of finite groups and associative algebras*, Pure and Applied Math., vol. XI, Interscience Publishers, New York-London, 1962.

[DO21] C. Daw & M. Orr – "Unlikely intersections with $E \times \mathrm{CM}$ curves in $\mathcal{A}_2$", *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (5)* **22** (2021), no. 4, p. 1705–1745.

[DO22] ———, "Quantitative reduction theory and unlikely intersections", *Internat. Math. Res. Notices* (2022), no. 20, p. 16138–16195.

[DR18] C. Daw & J. Ren – "Applications of the hyperbolic Ax–Schanuel conjecture", *Compositio Math.* **154** (2018), p. 1843–1888.

[DCD00] I. Del Corso & R. Dvornicich – "Relations among discriminant, different, and conductor of an order", *J. Algebra* **224** (2000), no. 1, p. 77–90.

[HP12] P. Habegger & J. Pila – "Some unlikely intersections beyond André–Oort", *Compositio Math.* **148** (2012), no. 1, p. 1–27.

[HP16] ———, "O-minimality and certain atypical intersections", *Ann. Sci. École Norm. Sup. (4)* **49** (2016), no. 4, p. 813–858.

[Hal35] P. Hall – "On representatives of subsets", *J. London Math. Soc.* **10** (1935), no. 1, p. 26–30.

[HJ85] R. A. Horn & C. R. Johnson – *Matrix analysis*, Cambridge University Press, Cambridge, 1985.

[KUY16] B. Klingler, E. Ullmo & A. Yafaev – "The hyperbolic Ax-Lindemann-Weierstrass conjecture", *Publ. Math. Inst. Hautes Études Sci.* **123** (2016), no. 1, p. 333–360.

[MW94] D. W. Masser & G. Wüstholz – "Endomorphism estimates for abelian varieties", *Math. Z.* **215** (1994), no. 4, p. 641–653.

[MW95] D. Masser & G. Wüstholz – "Factorization estimates for abelian varieties", *Publ. Math. Inst. Hautes Études Sci.* **81** (1995), p. 5–24.

[Mil99] J. S. Milne – "Lefschetz classes on abelian varieties", *Duke Math. J.* **96** (1999), no. 3, p. 639–675.

[Mil05] ———, "Introduction to Shimura varieties", in *Harmonic analysis, the trace formula, and Shimura varieties*, Clay Math. Proc., vol. 4, American Mathematical Society, Providence, RI, 2005, p. 265–378.

[Mum74] D. Mumford – *Abelian varieties*, second ed., T.I.F.R. Studies in Math., vol. 5, Oxford University Press, 1974.

[Orr15] M. Orr – "Families of abelian varieties with many isogenous fibres", *J. reine angew. Math.* **705** (2015), p. 211–231.

[Orr18] ———, "Height bounds and the Siegel property", *Algebra Number Theory* **12** (2018), no. 2, p. 455–478.

[Orr21] ———, "Unlikely intersections with Hecke translates of a special subvariety", *J. Eur. Math. Soc. (JEMS)* **23** (2021), no. 1, p. 1–28.

[PS10] Y. Peterzil & S. Starchenko – "Tame complex analysis and o-minimality", in *Proceedings of the ICM. Vol. II*, Hindustan Book Agency, New Delhi, 2010, p. 58–81.

[Pin05] R. Pink – "A common generalization of the conjectures of André–Oort, Manin–Mumford, and Mordell–Lang", preprint, available at `http://www.math.ethz.ch/~pink/ftp/AOMMML.pdf`, 2005.

[Rei75] I. Reiner – *Maximal orders*, London Math. Soc. Monographs, vol. 5, Academic Press, London-New York, 1975.

[Shi63] G. Shimura – "On analytic families of polarized abelian varieties and automorphic functions", *Ann. of Math. (2)* **78** (1963), p. 149–192.

[Ull14] E. Ullmo – "Applications du théorème d'Ax-Lindemann hyperbolique", *Compositio Math.* **150** (2014), no. 2, p. 175–190.

[Voi21] J. Voight – *Quaternion algebras*, Graduate Texts in Math., vol. 288, Springer, Cham, 2021.
[Wey40] H. Weyl – "Theory of reduction for arithmetical equivalence", *Trans. Amer. Math. Soc.* **48** (1940), p. 126–164.

Manuscript received 25th November 2021
accepted 10th May 2023

Christopher Daw, Department of Mathematics and Statistics, University of Reading
Whiteknights, PO Box 217, Reading, Berkshire RG6 6AH, United Kingdom
E-mail : chris.daw@reading.ac.uk
Url : https://sites.google.com/view/chrisdaw/home

Martin Orr, Department of Mathematics, The University of Manchester
Alan Turing Building, Oxford Road, Manchester M13 9PL, United Kingdom
E-mail : martin.orr@manchester.ac.uk
Url : https://personalpages.manchester.ac.uk/staff/Martin.Orr/