Manuscript version: Author's Accepted Manuscript

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:

http://wrap.warwick.ac.uk/142122

How to cite:

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

UNLIKELY INTERSECTIONS WITH $E \times \mathrm{CM}$ CURVES IN $\mathcal{A}_2$

CHRISTOPHER DAW AND MARTIN ORR

ABSTRACT. The Zilber–Pink conjecture predicts that an algebraic curve in $\mathcal{A}_2$ has only finitely many intersections with the special curves, unless it is contained in a proper special subvariety. Under a large Galois orbits conjecture, we prove the finiteness of the intersection with the special curves parametrising abelian surfaces isogenous to the product of two elliptic curves, at least one of which has complex multiplication. Furthermore, we show that this large Galois orbits conjecture holds for curves satisfying a condition on their intersection with the boundary of the Baily–Borel compactification of $\mathcal{A}_2$.

More generally, we show that a Hodge generic curve in an arbitrary Shimura variety has only finitely many intersection points with the generic points of a Hecke–facteur family, again under a large Galois orbits conjecture.

1. INTRODUCTION

Let $\mathcal{A}_2$ denote the moduli space of principally polarised abelian surfaces over $\mathbb{C}$. Let $V \subset \mathcal{A}_2$ be an irreducible algebraic curve. The Zilber–Pink conjecture predicts that, if V is not contained in any proper special subvariety of $\mathcal{A}_2$, then it should have only finitely many intersections with the special curves of $\mathcal{A}_2$ (since $\dim \mathcal{A}_2 = 3 > 1 + 1$).

The special curves in $\mathcal{A}_2$ are of three types:

(1) curves parametrising abelian surfaces with quaternionic multiplication;
(2) curves parametrising abelian surfaces isogenous to the square of an elliptic curve;
(3) curves parametrising abelian surfaces isogenous to the product of two elliptic curves, at least one of which has complex multiplication (CM).

In this paper, we consider the special curves of type (3), which we refer to as $E \times \mathrm{CM}$ curves. We consider the special curves of type (1) and type (2) in a forthcoming paper [DO].

There is a general strategy for proving conjectures of Zilber–Pink type known as the Pila–Zannier strategy. This was expounded for arbitrary Shimura varieties in [DR18]. The upshot of this strategy is that problems of unlikely intersections can be resolved whenever two arithmetic ingredients can be obtained. The first, usually referred to as a large Galois orbits conjecture, concerns the sizes of Galois

orbits of points arising from unlikely intersections. This is well-known to be a very difficult problem in general. The second ingredient, which is less well-known, concerns the parametrisation of special subvarieties and has much to do with the reduction theory of arithmetic groups.

In this article, we make progress on both these arithmetic ingredients for unlikely intersections with $E \times$ CM curves. We prove that the large Galois orbits conjecture for intersections between a general curve $V \subset \mathcal{A}_2$ and $E \times$ CM curves holds whenever V is defined over $\overline{\mathbb{Q}}$ and the closure of V in the Baily–Borel compactification of $\mathcal{A}_2$ intersects the zero-dimensional stratum of the boundary. This geometric condition on V can also be understood in terms of abelian schemes (see Proposition 9.4). It should be noted that this condition is not generic. It is central to the theorem of André on which our proof of the large Galois orbits conjecture is based [And89, Ch. X, 4.2 (i)].

Furthermore, we solve the parametrisation problem for $E \times$ CM curves in $\mathcal{A}_2$ in full. Consequently, we prove that if V is a curve not contained in any proper special subvariety of $\mathcal{A}_2$ then it has only finitely many intersections with $E \times$ CM curves, provided either that V satisfies the condition on its intersection with the boundary or that the large Galois orbits conjecture (Conjecture 1.3) holds.

1.A. **Statement of main theorems.** Recall that the Baily–Borel (or minimal) compactification of $\mathcal{A}_2$ has a stratification $\mathcal{A}_2 \sqcup \mathcal{A}_1 \sqcup \mathcal{A}_0$ by Zariski locally closed subvarieties, where $\mathcal{A}_1$ is the moduli space of elliptic curves, and $\mathcal{A}_0$ is a point (see [FC90, Ch. IV, sec. 2]). Our main theorem is as follows.

Theorem 1.1. *Let Σ denote the set of points in $\mathcal{A}_2$ for which the associated principally polarised abelian surface is isogenous to a product of elliptic curves $E_1 \times E_2$ such that E_2 has complex multiplication and E_1 does not.*

Let V be an irreducible algebraic curve in $\mathcal{A}_2$ which is not contained in any proper special subvariety.

Suppose that V is defined over $\overline{\mathbb{Q}}$ and that the Zariski closure of V in the Baily–Borel compactification of $\mathcal{A}_2$ intersects the 0-dimensional stratum of the boundary of the compactification.

Then $V \cap \Sigma$ is finite.

More generally, without the geometric assumption on V, we have the following conditional result.

Theorem 1.2. *Let Σ be as in Theorem 1.1. Let V be an irreducible algebraic curve in $\mathcal{A}_2$ which is not contained in any proper special subvariety.*

Assume that V satisfies Conjecture 1.3.

Then $V \cap \Sigma$ is finite.

The following large Galois orbits conjecture, referred to in the above, is similar to a special case of [DR18, Conjecture 11.1] (which was itself inspired by [HP16, Conjecture 8.2]) except that we use a definition of complexity adapted to $E \times$ CM

curves in place of the more general definition used in [DR18]. For the definition of the complexity $\Delta(Z)$, see section 3.

Conjecture 1.3. *Let V and Σ be as in Theorem 1.2. Let L be a finitely generated subfield of $\mathbb{C}$ over which V is defined.*

There exist positive constants C_1 and C_2 such that, for all points $s \in V \cap \Sigma$, if we let Z denote the (unique) special curve containing s, then

$$\# \operatorname{Aut}(\mathbb{C}/L) \cdot s \geq C_1 \Delta(Z)^{C_2}.$$

The statement of Conjecture 1.3 relies on the fact that each point of Σ lies in a *unique* special curve. This is true because the intersection of two distinct special curves consists of special points and we have excluded these from Σ. In other words, the set Σ is not equal to the union of the $E \times$ CM curves, because we have removed the (countably many) points where the corresponding abelian surface is isogenous to a product of two CM elliptic curves.

Removing these special points is important for Conjecture 1.3, because such a point lies in the intersection of infinitely many $E \times$ CM curves. Therefore we cannot expect the Galois orbit of such a point to be bounded below by the complexities of each of the $E \times$ CM curves which contain it. Instead, the Galois orbit of such a point is controlled by its own complexity as a special subvariety, as proved independently by Tsimerman [Tsi12] and by Ullmo and Yafaev [UY15].

Furthermore, Pila and Tsimerman have used this bound for special points to prove the André–Oort conjecture for $\mathcal{A}_2$ [PT13], that is, a non-special curve V in $\mathcal{A}_2$ contains only finitely many special points. Hence, a curve V as in Theorem 1.2 has finitely many intersections with the $E \times$ CM curves if and only if it contains only finitely many points belonging to Σ.

1.B. **Alternative statements in terms of abelian schemes.** It is also possible to state Theorems 1.1 and 1.2 using the language of abelian schemes instead of curves in $\mathcal{A}_2$. In these statements, if V is an algebraic curve and $\mathfrak{A} \to V$ is an abelian scheme of relative dimension 2, we say that a point $s \in V$ is an **$E \times$ CM point** if the fibre $\mathfrak{A}_s$ is isogenous to a product of elliptic curves $E_1 \times E_2$ where E_2 has complex multiplication and E_1 does not.

Our proof of Theorem 1.1 passes through the following theorem (we prove at Proposition 9.5 that this theorem implies Theorem 1.1).

Theorem 1.4. *Let V be an irreducible algebraic curve over $\mathbb{C}$ and let $\mathfrak{A} \to V$ be a principally polarised non-isotrivial abelian scheme of relative dimension 2. Suppose that $\operatorname{End}(\mathfrak{A}_{\overline{\eta}}) = \mathbb{Z}$, where $\overline{\eta}$ denotes a geometric generic point of V.*

Suppose that V and $\mathfrak{A}$ are defined over $\overline{\mathbb{Q}}$.

Suppose that there exist a curve V', a semiabelian scheme $\mathfrak{A}' \to V'$ and an open immersion $\iota \colon V \to V'$ (all defined over $\overline{\mathbb{Q}}$) such that $\mathfrak{A} \cong \iota^* \mathfrak{A}'$ and there is some point $s_0 \in V'(\overline{\mathbb{Q}})$ for which the fibre $\mathfrak{A}'_{s_0}$ is a torus.

Then V contains only finitely many $E \times$ CM points.

Theorem 1.2 can be reformulated in a similar style to Theorem 1.4. Likewise, Conjecture 1.3 is equivalent to the following conjecture (see Proposition 9.1). The complexity $\Delta'(s)$ used in this conjecture is defined in Remark 3.3.

Conjecture 1.5. *Let V be an irreducible algebraic curve over $\mathbb{C}$ and let $\mathfrak{A} \to V$ be a principally polarised non-isotrivial abelian scheme of relative dimension 2. Suppose that $\mathrm{End}(\mathfrak{A}_{\overline{\eta}}) = \mathbb{Z}$, where $\overline{\eta}$ denotes a geometric generic point of V.*

Let L be a finitely generated subfield of $\mathbb{C}$ over which V and $\mathfrak{A} \to V$ are defined. There exist positive constants C_3 and C_4 such that, for all $E \times CM$ points $s \in V$,

$$\# \mathrm{Aut}(\mathbb{C}/L) \cdot s \geq C_3 \Delta'(s)^{C_4}.$$

1.C. **Previous results.** Of course, Unlikely Intersections is now a vast area of research. This article can be viewed as a sequel to two previous works. The first [DR18], due to Ren and the first author, established a general strategy to attack the Zilber–Pink conjecture for general Shimura varieties, depending on certain arithmetic conjectures. The second [Orr], due to the second author, proved both conditional and unconditional results for unlikely intersections with Hecke translates of a fixed special subvariety. Both of these works were inspired by the earlier works [HP12] and [HP16] of Habegger and Pila on the Zilber–Pink conjecture in a product of modular curves. In this article, we implement the stategy of [DR18] for so-called $E \times$ CM curves in $\mathcal{A}_2$ and, more generally, for so-called Hecke–facteur families, which are a natural generalisation of the objects studied in [Orr].

Considering special points instead of intersections with special curves, the André–Oort conjecture for $\mathcal{A}_2$ was proved by Pila and Tsimerman [PT13], also using the Pila–Zannier strategy. This strategy eventually led to their proof of the André–Oort conjecture for $\mathcal{A}_g$ [PT14], [Tsi18].

The key step in proving Theorem 1.2 consists in controlling the heights of algebraic points in definable sets which parametrise intersections with special subvarieties. Analogous bounds for pre-special points of $\mathcal{A}_g$ appeared in [PT13], and were generalised by the authors of the current paper to arbitrary Shimura varieties [DO16]. The second author has proved bounds of a similar nature concerning Hecke operators [Orr18]. We use both of these previous bounds in this paper.

1.D. **Generalisations.** We prove a generalisation of Theorem 1.2 using the notion of Hecke–facteur families in a general Shimura variety, which we define in section 2.D. These families are a natural generalisation of $E \times$ CM curves in $\mathcal{A}_2$.

For an example of Hecke–facteur families, fix positive integers d, e, g satisfying $d + e = g$. Consider the locus of points in $\mathcal{A}_g$ parametrising principally polarised abelian varieties which are isogenous (via a polarised isogeny) to a product $A_1 \times A_2$, where $\dim(A_1) = d$, $\dim(A_2) = e$ and A_2 has CM type. This locus is the union of a Hecke–facteur family of special subvarieties of $\mathcal{A}_g$, as described in section 2.E. The example includes the case of $E \times$ CM curves by choosing $d = e = 1$. (Note that the definition of $E \times$ CM curves does not require the isogeny to be polarised. Lemma 2.1 shows that this does not matter in the $E \times$ CM case.)

Another example of a Hecke–facteur family is given by the Hecke translates of a fixed special subvariety (this is obtained by setting $\mathbf{H}_2 = \{1\}$ in the definition of Hecke–facteur family). This case of Theorem 1.6 was essentially already established in [Orr, Theorem 3.4]. We are able to remove the dependence on [Orr, Conjecture 3.3] thanks to Proposition 4.3.

Theorem 1.6 is conditional on a Galois orbits conjecture, Conjecture 1.7. We do not prove any cases of Conjecture 1.7 beyond the $E \times \mathrm{CM}$ case.

Theorem 1.6. *Let $(\mathbf{G}, X^+)$ be a Shimura datum component and let $S = \Gamma \backslash X^+$ be an associated Shimura variety component. Let V be an irreducible algebraic curve in S not contained in any proper special subvariety.*

Let $(\mathbf{H}, X_{\mathbf{H}}^+, \mathbf{H}_1, \mathbf{H}_2)$ be a facteur datum in $(\mathbf{G}, X^+)$ and let $\mathfrak{F}$ be the associated Hecke–facteur family of special subvarieties of S. Suppose that, for every subvariety $Z \in \mathfrak{F}$, we have

$$\dim(Z) \leq \dim(S) - 2.$$

Let Σ be the set of points $s \in S$ for which the smallest special subvariety containing s is a member of $\mathfrak{F}$.

Assume that S, $\mathfrak{F}$, V and Σ satisfy Conjecture 1.7.

Then $V \cap \Sigma$ is finite.

The following large Galois orbits conjecture gives a lower bound for the Galois degree of a point $s \in V \cap \Sigma$ with respect to the complexity of the special subvariety of $\mathfrak{F}$ containing s. The complexity $\Delta(Z)$ of a special subvariety in the Hecke–facteur family $\mathfrak{F}$ is defined in section 3, depending on the choice of a representation ρ (hence why we need to make such a choice in the statement of the conjecture).

Conjecture 1.7. *Let $(\mathbf{G}, X^+)$, S, $\mathfrak{F}$, V and Σ be as in Theorem 1.6. Let L be a finitely generated subfield of $\mathbb{C}$ over which S and V are defined. Let $\rho\colon \mathbf{G} \to \mathbf{GL}_{m,\mathbb{Q}}$ be a faithful representation.*

There exist positive constants C_5 and C_6 such that, for all points $s \in V \cap \Sigma$, if we let Z denote the (unique) special subvariety in $\mathfrak{F}$ containing s, then

$$\# \operatorname{Aut}(\mathbb{C}/L) \cdot s \geq C_5 \Delta(Z)^{C_6}.$$

1.E. **Strategy.** The strategy to prove Theorem 1.2 is to prove [DR18, Conjecture 12.2] for the pre-images of the points in Σ (modified to use a different definition of complexity). This is achieved by proving [DR18, Conjecture 12.7] for these points, which is deduced from a new height bound for representatives in congruence subgroups (Proposition 5.1), and combining it with previous results of the authors ([DO16, Theorems 1.1 and 4.1] and [Orr18, Theorem 1.1]). Theorem 1.2 then follows by the method of [DR18, Theorem 14.2] (note that [DR18, Conjecture 10.3] follows from [DR18, Conjecture 12.2] by [DR18, Lemma 12.5]).

Strictly speaking, we cannot directly invoke [DR18, Theorem 14.2]; rather, we have to show that its proof applies to the points in Σ. This requires that the $E \times \mathrm{CM}$ curves are permuted by a Galois action, which is ensured by Lemma 4.1.

The above arguments work for any Hecke–facteur family in a Shimura variety, thus proving Theorem 1.6. Specialising to the $E \times$ CM case, the proof of Theorems 1.1 and 1.4 is obtained by adapting a height bound of André for abelian varieties with large endomorphism rings [And89, Ch. X, Theorem 1.3], combined with the Masser–Wüstholz isogeny theorem.

1.F. **Outline of paper.** Section 2 contains various definitions, notation and basic facts relating to Shimura varieties, Hecke–facteur families and Siegel fundamental sets. Section 3 is largely devoted to comparing the two definitions of complexity used for Conjectures 1.3 and 1.5 respectively.

In section 4, we prove that a suitable Galois action permutes the $E \times$ CM curves, as required to be able to apply the method of [DR18, Theorem 14.2]. Section 5 proves a height bound for congruence subgroups which is another essential ingredient in the proof of Theorem 1.2. We carry out the strategy of [DR18] for Hecke–facteur families in Sections 6 and 7. Section 6 establishes height bounds for parameters of special subvarieties (modified versions of [DR18, Conjectures 12.2 and 12.7]) while Section 7 describes how we need to modify the proof of [DR18, Theorem 14.2] to obtain Theorems 1.2 and 1.6.

Finally Sections 8 and 9 concern Galois orbits bounds for intersections with $E \times$ CM curves. Section 8 generalises the height bound of [And89, Ch. X] to include abelian surfaces. Section 9 applies this to prove Theorems 1.1 and 1.4.

2. Preliminaries

2.A. **Shimura varieties.** Let $\mathbb{S}$ denote the Deligne torus $\mathrm{Res}_{\mathbb{C}/\mathbb{R}} \mathbb{G}_m$. A **Shimura datum** is a pair $(\mathbf{G}, X)$, where $\mathbf{G}$ is a connected reductive $\mathbb{Q}$-algebraic group and X is a $\mathbf{G}(\mathbb{R})$-conjugacy class in $\mathrm{Hom}(\mathbb{S}, \mathbf{G}_{\mathbb{R}})$ satisfying [Del79, axioms 2.1.1.1–2.1.1.3]. These axioms imply that X is a finite disjoint union of Hermitian symmetric domains [Del79, Corollaire 1.1.17].

Given a Shimura datum $(\mathbf{G}, X)$ and a compact open subgroup $K \subset \mathbf{G}(\mathbb{A}_f)$, the resulting **Shimura variety** is denoted $\mathrm{Sh}_K(\mathbf{G}, X)$. This is a quasi-projective algebraic variety whose complex points are given by

$$\mathrm{Sh}_K(\mathbf{G}, X)(\mathbb{C}) = \mathbf{G}(\mathbb{Q}) \backslash X \times \mathbf{G}(\mathbb{A}_f)/K.$$

We can attach to each Shimura datum a number field called the **reflex field**, denoted $E(\mathbf{G}, X)$. According to Deligne's theory, $\mathrm{Sh}_K(\mathbf{G}, X)$ has a so-called **canonical model** over $E(\mathbf{G}, X)$ (see [Del79], completed in [MS82], [Mil83] and [Bor84]).

Given a morphism of Shimura data $(\mathbf{H}, X_{\mathbf{H}}) \to (\mathbf{G}, X)$, induced from a morphism $f : \mathbf{H} \to \mathbf{G}$ of $\mathbb{Q}$-algebraic groups, and compact open subgroups $K_{\mathbf{H}} \subset \mathbf{H}(\mathbb{A}_f)$ and $K \subset \mathbf{G}(\mathbb{A}_f)$ such that $f(K_{\mathbf{H}}) \subset K$, we obtain a morphism of Shimura varieties

$$\mathrm{Sh}_{K_{\mathbf{H}}}(\mathbf{H}, X_{\mathbf{H}}) \to \mathrm{Sh}_K(\mathbf{G}, X),$$

which is closed and defined over the compositum of the reflex fields $E_{\mathbf{H}} := E(\mathbf{H}, X_{\mathbf{H}})$ and $E_{\mathbf{G}} := E(\mathbf{G}, X_{\mathbf{G}})$ [Mil05, Remark 13.8]. If f is an inclusion, then $E_{\mathbf{G}} \subset E_{\mathbf{H}}$.

Let $\mathbf{H}^{\mathrm{ad}}$ denote the quotient of $\mathbf{H}$ by its centre, and let $f : \mathbf{H} \to \mathbf{H}^{\mathrm{ad}}$ be the natural morphism. Then we obtain a Shimura datum $(\mathbf{H}^{\mathrm{ad}}, X_{\mathbf{H}}^{\mathrm{ad}})$ where $X_{\mathbf{H}}^{\mathrm{ad}}$ is the $\mathbf{H}^{\mathrm{ad}}(\mathbb{R})$-conjugacy class of morphisms containing the image of $X_{\mathbf{H}}$ under composition with $f_{\mathbb{R}}$. The reflex fields are related by $E(\mathbf{H}^{\mathrm{ad}}, X_{\mathbf{H}}^{\mathrm{ad}}) \subset E_{\mathbf{H}}$.

A key example of a Shimura variety is $\mathcal{A}_g$, the coarse moduli space of principally polarised abelian varieties of dimension g. This Shimura variety is geometrically irreducible and defined over $\mathbb{Q}$. It is equal to $\mathrm{Sh}_K(\mathbf{G}, X)$ where $\mathbf{G}$ is the general symplectic group $\mathbf{GSp}_{2g}$, X is isomorphic to the disjoint union of two copies of the Siegel upper half-space $\mathcal{H}_g$ and $K = \mathbf{GSp}_{2g}(\hat{\mathbb{Z}})$.

For any algebraically closed field k and any point $s \in \mathcal{A}_g(k)$, we shall write A_s for the abelian variety parametrised by s (which is defined up to k-isomorphism).

2.B. **Shimura variety components.** Over $\mathbb{C}$, a Shimura variety usually has many irreducible components. It will often be convenient for us to work with a single geometrically irreducible component of a Shimura variety, which we call a **Shimura variety component**.

We define a **Shimura datum component** to be a pair $(\mathbf{G}, X^+)$, where $(\mathbf{G}, X)$ is a Shimura datum and X^+ is a connected component of X. Let $\mathbf{G}(\mathbb{Q})_+$ denote the stabiliser of X^+ in $\mathbf{G}(\mathbb{Q})$. The image of X^+ in $\mathrm{Sh}_K(\mathbf{G}, X)$ is a Shimura variety component, which we will denote by $\mathrm{Sh}_K(\mathbf{G}, X^+)$. The complex points of $\mathrm{Sh}_K(\mathbf{G}, X^+)$ are isomorphic (as a complex analytic space) to $\Gamma \backslash X^+$ where $\Gamma = K \cap \mathbf{G}(\mathbb{Q})_+$ (subgroups of $\mathbf{G}(\mathbb{Q})_+$ of this form are called **congruence subgroups**).

We say that $(\mathbf{H}, X_{\mathbf{H}}^+)$ is a **Shimura subdatum component** of $(\mathbf{G}, X^+)$ if $(\mathbf{G}, X^+)$ and $(\mathbf{H}, X_{\mathbf{H}}^+)$ are Shimura datum components, $\mathbf{H}$ is an algebraic subgroup of $\mathbf{G}$ and the inclusion $\mathbf{H} \hookrightarrow \mathbf{G}$ induces an inclusion $X_{\mathbf{H}}^+ \hookrightarrow X^+$.

Let $(\mathbf{H}, X_{\mathbf{H}}^+)$ be a Shimura subdatum component of $(\mathbf{G}, X^+)$. Let $\Gamma \subset \mathbf{G}(\mathbb{Q})_+$ be a congruence subgroup and let $\Gamma_{\mathbf{H}} = \Gamma \cap \mathbf{H}(\mathbb{Q})_+$. Then $X_{\mathbf{H}}^+ \hookrightarrow X^+$ induces a closed morphism $\Gamma_{\mathbf{H}} \backslash X_{\mathbf{H}}^+ \to \Gamma \backslash X^+$ between Shimura variety components. We call the image of $X_{\mathbf{H}}^+ \to X^+$ a **pre-special subvariety** of X^+ and the image of $\Gamma_{\mathbf{H}} \backslash X_{\mathbf{H}}^+ \to \Gamma \backslash X^+$ a **special subvariety** of $\Gamma \backslash X^+$.

2.C. **Hecke correspondences.** Let $(\mathbf{G}, X)$ be a Shimura datum and let $K \subset \mathbf{G}(\mathbb{A}_f)$ be a compact open subgroup. For each $g \in \mathbf{G}(\mathbb{A}_f)$, there is an algebraic correspondence on $\mathrm{Sh}_K(\mathbf{G}, X)$ given by the morphisms

$$\mathrm{Sh}_K(\mathbf{G}, X) \leftarrow \mathrm{Sh}_{K \cap gKg^{-1}}(\mathbf{G}, X) \to \mathrm{Sh}_{g^{-1}Kg \cap K}(\mathbf{G}, X) \to \mathrm{Sh}_K(\mathbf{G}, X),$$

where the middle arrow is induced by the map

$$(x, a) \to (x, ag) : X \times \mathbf{G}(\mathbb{A}_f) \to X \times \mathbf{G}(\mathbb{A}_f),$$

and the outside arrows are the natural projections. We call such a correspondence a **Hecke correspondence** and denote by T_g the induced map on algebraic cycles.

2.D. **Hecke–facteur families.** Let $(\mathbf{G}, X^+)$ be a Shimura datum component and let $(\mathbf{H}, X_{\mathbf{H}}^+) \subset (\mathbf{G}, X^+)$ be a Shimura subdatum component.

Suppose that $\mathbf{H}$ has semisimple normal subgroups $\mathbf{H}_1$ and $\mathbf{H}_2$ such that $\mathbf{H}_1 \cap \mathbf{H}_2$ is finite and $\mathbf{H}^{\mathrm{der}} = \mathbf{H}_1.\mathbf{H}_2$. The groups $\mathbf{H}_1$ and $\mathbf{H}_2$ are not necessarily associated with Shimura data, but their adjoint groups give rise to Shimura datum components $(\mathbf{H}_1^{\mathrm{ad}}, X_1^+)$ and $(\mathbf{H}_2^{\mathrm{ad}}, X_2^+)$ such that there is a natural isomorphism $X_{\mathbf{H}}^+ \cong X_1^+ \times X_2^+$. We henceforth regard this isomorphism as an *identification* and write $X_1^+ \times X_2^+ = X_{\mathbf{H}}^+ \subset X^+$ without extra notation. (In the language of [Mil05, Definition 4.4], $(\mathbf{H}_1, X_1^+)$ and $(\mathbf{H}_2, X_2^+)$ are connected Shimura data.)

We call $(\mathbf{H}, X_{\mathbf{H}}^+, \mathbf{H}_1, \mathbf{H}_2)$ as above a **facteur datum** in $(\mathbf{G}, X^+)$.

Let $(\mathbf{H}, X_{\mathbf{H}})$ denote the Shimura datum such that $X_{\mathbf{H}}^+$ is a component of $X_{\mathbf{H}}$. The decomposition $X_{\mathbf{H}}^+ \cong X_1^+ \times X_2^+$ extends to a decomposition of Shimura data $(\mathbf{H}^{\mathrm{ad}}, X_{\mathbf{H}}^{\mathrm{ad}}) \cong (\mathbf{H}_1, X_1) \times (\mathbf{H}_2, X_2)$. The reflex field $E(\mathbf{H}^{\mathrm{ad}}, X_{\mathbf{H}}^{\mathrm{ad}})$ is equal to the compositum of $E(\mathbf{H}_1, X_1)$ and $E(\mathbf{H}_2, X_2)$. If $K_{\mathbf{H}}^{\mathrm{ad}} \subset \mathbf{H}^{\mathrm{ad}}(\mathbb{A}_f)$ is a compact open subgroup equal to a product of compact open subgroups $K_1 \subset \mathbf{H}_1(\mathbb{A}_f)$ and $K_2 \subset \mathbf{H}_2(\mathbb{A}_f)$, then we obtain a decomposition of Shimura varieties

$$\mathrm{Sh}_{K_{\mathbf{H}}^{\mathrm{ad}}}(\mathbf{H}^{\mathrm{ad}}, X_{\mathbf{H}}^{\mathrm{ad}}) = \mathrm{Sh}_{K_1}(\mathbf{H}_1^{\mathrm{ad}}, X_1) \times \mathrm{Sh}_{K_2}(\mathbf{H}_2^{\mathrm{ad}}, X_2),$$

which is defined over $E(\mathbf{H}^{\mathrm{ad}}, X_{\mathbf{H}}^{\mathrm{ad}})$.

For any pre-special point $x_2 \in X_2^+$, $X_1^+ \times \{x_2\}$ is a pre-special subvariety of X^+. We call the collection of subvarieties of this form (for a fixed facteur datum) a **facteur family of pre-special subvarieties** of X^+. We say that a facteur family is **trivial** if it comes from a facteur datum with $\mathbf{H}_2 = \{1\}$: in this case, the facteur family consists simply of $X_{\mathbf{H}}^+$ itself. This terminology is motivated by the notion of "non-facteur" special subvarieties from [Ull07]: a special subvariety is "non-facteur" if and only if it does not belong to any non-trivial facteur family.

The main topic of this paper will be the family of Hecke translates of a given facteur family. For any $g \in \mathbf{G}(\mathbb{Q})_+$ and any pre-special point $x_2 \in X_2^+$,

$$Y_{g, x_2} = g(X_1^+ \times \{x_2\})$$

is again a pre-special subvariety of X^+. We call the collection of subvarieties Y_{g, x_2} a **Hecke–facteur family of pre-special subvarieties** of X^+. We call their images $Z_{g, x_2} = \pi(Y_{g, x_2})$, where π denotes the uniformisation map $X^+ \to \Gamma \backslash X^+$, a **Hecke–facteur family of special subvarieties** of $\Gamma \backslash X^+$. We chose this name because Z_{g, x_2} is an irreducible component of the image of $\pi(X_1^+ \times \{x_2\})$ (a special subvariety in the facteur family) under the Hecke correspondence $T_{g^{-1}}$.

The Mumford–Tate group $\mathrm{MT}(Y_{g,x_2})$ of a very general point in Y_{g,x_2} is equal to $g(\mathbf{H}_1.\mathbf{T}_{x_2})g^{-1}$, where $\mathbf{T}_{x_2}$ is a torus in $Z(\mathbf{H}).\mathbf{H}_2$ corresponding to the pre-special point $x_2 \in X_2^+$. Observe that $\mathrm{MT}(Y_{g,x_2})^{\mathrm{der}} = g\mathbf{H}_1 g^{-1}$.

2.E. **Example of a Hecke–facteur family.** For a fundamental example of a Hecke–facteur family, choose positive integers d, e, g such that $d + e = g$. Let $(\mathbf{G}, X^+) = (\mathbf{GSp}_{2g}, \mathcal{H}_g)$ and let $X_{\mathbf{H}}^+ \subset X^+$ be a pre-special subvariety parametrising principally polarised abelian varieties which are isomorphic (as polarised abelian varieties) to a product $A \times B$, where $\dim(A) = d$, $\dim(B) = e$ and A, B are principally polarised. The generic Mumford–Tate group of $X_{\mathbf{H}}^+$ is

$$\mathbf{H} = \mathbb{G}_m.(\mathbf{Sp}_{2d} \times \mathbf{Sp}_{2e}).$$

Let $\mathbf{H}_1 = \mathbf{Sp}_{2d} \subset \mathbf{H}$ and $\mathbf{H}_2 = \mathbf{Sp}_{2e} \subset \mathbf{H}$. Then $(\mathbf{H}, X_{\mathbf{H}}^+, \mathbf{H}_1, \mathbf{H}_2)$ is a facteur datum in $(\mathbf{G}, X^+)$, and $X_{\mathbf{H}}^+ = X_1^+ \times X_2^+$, where $X_1^+ = \mathcal{H}_d$ and $X_2^+ = \mathcal{H}_e$. The resulting Hecke–facteur family consists of the special subvarieties which parametrise principally polarised abelian varieties A of dimension g such that there exist principally polarised abelian varieties A_1 (of dimension d) and A_2 (of dimension e and of CM type) and a polarised isogeny $A_1 \times A_2 \to A$. (Given polarised abelian varieties (A, λ) and (B, μ), a **polarised isogeny** is an isogeny $f \colon A \to B$ such that $f^*\mu = n\lambda$ for some $n \in \mathbb{Z}$.)

In this paper, we are principally interested in the case $g = 2$, $d = e = 1$ of this construction. The following lemma shows that, in this case, the special curves in the Hecke–facteur family contain every point of $\mathcal{A}_2$ for which the corresponding abelian variety is isogenous (not just *polarised* isogenous) to a product $E_1 \times E_2$ where E_2 has CM. In other words, this Hecke–facteur family consists precisely of the $E \times$ CM curves defined in the introduction.

Lemma 2.1. *Let (A, λ) be a principally polarised abelian surface for which there exist elliptic curves E_1 and E_2 (not in the same isogeny class) and an isogeny $\varphi \colon E_1 \times E_2 \to A$.*

Then there exist elliptic curves E_1' and E_2', such that E_i' is isogenous to E_i, and a polarised isogeny $\varphi' \colon E_1' \times E_2' \to (A, \lambda)$ (where E_1', E_2' are equipped with their principal polarisations).

Furthermore, $\deg \varphi' \leq \deg \varphi$.

Proof. Let $\Lambda = H_1(A, \mathbb{Z})$, equipped with the symplectic pairing ψ induced by the polarisation λ. The isogeny $\varphi \colon E_1 \times E_2 \to A$ induces an injection of $\mathbb{Z}$-modules

$$H_1(E_1, \mathbb{Z}) \oplus H_1(E_2, \mathbb{Z}) \to \Lambda$$

with finite cokernel.

Let $\Lambda_i = (H_1(E_i, \mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{Q}) \cap \Lambda$. This is the $\mathbb{Z}$-Hodge structure of an elliptic curve E_i' isogenous to E_i. The inclusion $\Lambda_1 \oplus \Lambda_2 \to \Lambda$ induces an isogeny $\varphi' \colon E_1' \times E_2' \to A$.

Let λ_i' be the (unique) principal polarisation on E_i' and let ψ_i' be the associated symplectic form on Λ_i. Because $\mathrm{rk}\,\Lambda_i = 2$, every symplectic form on Λ_i is an integer

multiple of ψ_i'. (This is the step in the proof which is restricted to a product of elliptic curves.) In particular $\psi_{|\Lambda_i} = n_i\psi_i'$ for some $n_i \in \mathbb{Z}$. Because ψ and ψ_i' can both be interpreted as the imaginary part of a positive definite Hermitian form (because they come from polarisations), we must have $n_i > 0$.

Because E_1 and E_2 are non-isogenous, there is no non-zero morphism of $\mathbb{Z}$-Hodge structures $\Lambda_1 \to \Lambda_2$ or $\Lambda_2 \to \Lambda_1$. Therefore Λ_1 and Λ_2 are orthogonal with respect to ψ. It follows that $\psi = n_1\psi_1' + n_2\psi_2'$ and so $\varphi'^*\lambda = (n_1\lambda_1', n_2\lambda_2')$.

By construction Λ_1 and Λ_2 are primitive submodules of Λ. Since $\mathrm{disc}(n_i\psi_i') = n_i^2$, Lemma 2.2 implies that $n_1 = n_2$ so $\varphi'^*\lambda$ is an integer multiple of (λ_1', λ_2'), that is, φ' is a polarised isogeny.

Furthermore,

$$H_1(E_1, \mathbb{Z}) \oplus H_1(E_2, \mathbb{Z}) \subset \Lambda_1 \oplus \Lambda_2 \subset \Lambda.$$

It follows that φ factors through φ' and so $\deg \varphi' \leq \deg \varphi$. $\square$

In the proof of Lemma 2.1, we needed the following lemma. This is a symplectic version of [Huy16, Chapter 14, Proposition 0.2], which is the analogous result for symmetric bilinear forms.

Lemma 2.2. *Let Λ be a free $\mathbb{Z}$-module of finite rank with a perfect symplectic form $\psi \colon \Lambda \times \Lambda \to \mathbb{Z}$. Let Λ_1, Λ_2 be primitive submodules of Λ which are orthogonal to each other and such that $\Lambda_1 \oplus \Lambda_2$ has finite index in Λ. Then*

$$\mathrm{disc}(\psi_{|\Lambda_1}) = \mathrm{disc}(\psi_{|\Lambda_2}) = [\Lambda : \Lambda_1 + \Lambda_2].$$

Proof. Let

$$\Lambda_i^\vee = \{v \in \Lambda_i \otimes \mathbb{Q} : \psi(v, \Lambda_i) \subset \mathbb{Z}\}.$$

The map $v \mapsto \psi_{|\Lambda_i}(v, -)$ is an isomorphism $\Lambda_i^\vee \to \mathrm{Hom}(\Lambda_i, \mathbb{Z})$. Consequently, there is a unique $\mathbb{Z}$-module homomorphism $\alpha_i \colon \Lambda \to \Lambda_i^\vee$ such that

$$\psi_{|\Lambda_i}(v, -) = \psi_{|\Lambda_i}(\alpha_i(v), -)$$

for all $v \in \Lambda$.

We shall show that α_i induces an isomorphism $\Lambda/(\Lambda_1 + \Lambda_2) \to \Lambda_i^\vee/\Lambda_i$ for each i. This will suffice to prove the lemma because $\mathrm{disc}(\psi_{|\Lambda_i}) = [\Lambda_i^\vee : \Lambda_i]$.

Firstly, $\alpha_i \colon \Lambda \to \Lambda_i^\vee$ is surjective because Λ_i is a primitive submodule of Λ and ψ is a perfect pairing on Λ.

For every $v \in \Lambda$, $v - \alpha_1(v)$ is orthogonal to Λ_1. Thus $v - \alpha_1(v) \in \Lambda_2 \otimes \mathbb{Q}$.

If $\alpha_1(v) \in \Lambda_1$, then $v - \alpha_1(v) \in \Lambda$. Since $v - \alpha_1(v) \in \Lambda_2 \otimes \mathbb{Q}$ and Λ_2 is a primitive submodule of Λ, we deduce that $v - \alpha_1(v) \in \Lambda_2$ and hence $v \in \Lambda_1 + \Lambda_2$. Conversely, if $v \in \Lambda_1 + \Lambda_2$, then writing $v = v_1 + v_2$ with $v_i \in \Lambda_i$, we get $\alpha_1(v) = v_1 \in \Lambda_1$.

Thus $\alpha_1(v) \in \Lambda_1$ if and only if $v \in \Lambda_1 + \Lambda_2$. In other words, α_1 induces an injection $\Lambda/(\Lambda_1 + \Lambda_2) \to \Lambda_1^\vee/\Lambda_1$. A similar argument applies to α_2. $\square$

TABLE 1. Special subvarieties of $\mathcal{A}_2$

Description	Dimension	Generic endomorphism algebra
$\mathcal{A}_2$	3	$\mathbb{Q}$
Hilbert modular surface	2	real quadratic field
Hecke translate of $\mathcal{A}_1 \times \mathcal{A}_1$	2	$\mathbb{Q} \times \mathbb{Q}$
Shimura curve	1	indefinite quaternion algebra over $\mathbb{Q}$
modular curve	1	$\mathrm{M}_2(\mathbb{Q})$
$E \times \mathrm{CM}$ curve	1	$\mathbb{Q} \times F$, where F is an imaginary quadratic field
simple CM point	0	quartic CM field
non-simple CM point	0	product of two distinct imaginary quadratic fields
isotypic CM point	0	$\mathrm{M}_2(F)$, where F is an imaginary quadratic field

2.F. **Special subvarieties of $\mathcal{A}_2$.** The main theorems of this paper concern special subvarieties of $\mathcal{A}_2$. As a consequence of the classification of Hodge groups of abelian surfaces in [MZ99], every special subvariety of $\mathcal{A}_2$ is an irreducible component of the locus of principally polarised abelian surfaces whose endomorphism algebras contain a fixed algebra ("a Shimura variety of PEL type").

In Table 1, we list the classes of special subvarieties of $\mathcal{A}_2$, giving the dimension of each special subvariety and the generic endomorphism algebra of the abelian surfaces parametrised by that special subvariety.

Unlikely intersections with special points are taken care of by the André–Oort conjecture (proved for $\mathcal{A}_2$ by Pila and Tsimerman [PT13]). Therefore, in order to prove the Zilber–Pink conjecture for $\mathcal{A}_2$, it remains only to consider intersections between special curves and a general curve. In this paper, we consider intersections with the $E \times \mathrm{CM}$ special curves. The outstanding special curves will be treated in a forthcoming article by the same authors [DO].

2.G. **Fundamental sets.** Let $(\mathbf{G}, X^+)$ be a Shimura datum component and let $(\mathbf{H}, X_\mathbf{H}^+, \mathbf{H}_1, \mathbf{H}_2)$ be a facteur datum. Let $(\mathbf{H}_1^{\mathrm{ad}}, X_1^+)$ and $(\mathbf{H}_2^{\mathrm{ad}}, X_2^+)$ denote the Shimura datum components such that $X_\mathbf{H}^+ \cong X_1^+ \times X_2^+$. We now describe how to choose compatible fundamental sets in X_1^+, X_2^+ and X^+.

Definition. A **Siegel fundamental set** in X^+ for a congruence subgroup $\Gamma \subset \mathbf{G}(\mathbb{Q})_+$ is a fundamental set for Γ of the form $C.\mathfrak{S}^+.x_0$, where $C \subset \mathbf{G}(\mathbb{Q})_+$ is a finite set, $\mathfrak{S}^+ = \mathfrak{S} \cap \mathbf{G}(\mathbb{R})^+$ for some Siegel set $\mathfrak{S} \subset \mathbf{G}(\mathbb{R})$ and $x_0 \in X^+$ is a point

such that the stabiliser of x_0 in $\mathbf{G}(\mathbb{R})$ right-stabilises $\mathfrak{S}$. We use the definitions of Siegel sets and associated terminology from [Orr18, section 2B].

Lemma 2.3. *Fix a congruence subgroup $\Gamma \subset \mathbf{G}(\mathbb{Q})_+$ and let $\Gamma_1 = \Gamma \cap \mathbf{H}_1(\mathbb{Q})_+$, $\Gamma_2 = \Gamma \cap \mathbf{H}_2(\mathbb{Q})_+$. Let $\mathcal{F}_1 \subset X_1^+$ and $\mathcal{F}_2 \subset X_2^+$ be Siegel fundamental sets for the congruence subgroups Γ_1 and Γ_2, respectively.*

Then there exists a Siegel fundamental set $\mathcal{F} \subset X^+$ for Γ such that $\mathcal{F}_1 \times \mathcal{F}_2 \subset \mathcal{F}$.

The proof of Lemma 2.3 relies on the following lemma.

Lemma 2.4. *Let $\mathfrak{S}_1 \subset \mathbf{H}_1(\mathbb{R})$ and $\mathfrak{S}_2 \subset \mathbf{H}_2(\mathbb{R})$ be Siegel sets. Then $\mathfrak{S}_1.\mathfrak{S}_2$ is a Siegel set in $\mathbf{H}^{\mathrm{der}}(\mathbb{R})$.*

Proof. We use the notation from [Orr18, section 2B], adding subscripts 1, 2 or $\mathbf{H}^{\mathrm{der}}$ as appropriate. For example, $(\mathbf{P}_1, \mathbf{S}_1, K_1)$ denotes the Siegel triple associated with the Siegel set $\mathfrak{S}_1$.

We begin by constructing a Siegel triple for $\mathbf{H}^{\mathrm{der}}$. Multiplication in $\mathbf{H}$ is a central $\mathbb{Q}$-isogeny $\mathbf{H}_1 \times \mathbf{H}_2 \to \mathbf{H}^{\mathrm{der}}$, so $\mathbf{P}_{\mathbf{H}^{\mathrm{der}}} = \mathbf{P}_1.\mathbf{P}_2$ is a minimal parabolic $\mathbb{Q}$-subgroup of $\mathbf{H}^{\mathrm{der}}$. Similarly $K_{\mathbf{H}^{\mathrm{der}}} = K_1.K_2$ is a maximal compact subgroup of $\mathbf{H}^{\mathrm{der}}(\mathbb{R})$.

Let $\mathbf{S}_{\mathbf{H}^{\mathrm{der}}} = \mathbf{S}_1.\mathbf{S}_2$. This is an $\mathbb{R}$-torus in $\mathbf{P}_{\mathbf{H}^{\mathrm{der}}}$. Since $\mathbf{S}_i$ is $\mathbf{P}_i(\mathbb{R})$-conjugate to a maximal $\mathbb{Q}$-split torus in $\mathbf{H}_i$, we can use the central $\mathbb{Q}$-isogeny to deduce that $\mathbf{S}_{\mathbf{H}^{\mathrm{der}}}$ is $\mathbf{P}_{\mathbf{H}^{\mathrm{der}}}(\mathbb{R})$-conjugate to a maximal $\mathbb{Q}$-split torus in $\mathbf{P}_{\mathbf{H}^{\mathrm{der}}}$. Finally $\mathbf{S}_{\mathbf{H}^{\mathrm{der}}}$ is stabilised by the Cartan involution of $\mathbf{H}^{\mathrm{der}}$ associated with $K_{\mathbf{H}^{\mathrm{der}}}$, because this Cartan involution restricts to the Cartan involutions of $\mathbf{H}_1$ and $\mathbf{H}_2$ associated with K_1 and K_2 respectively. Thus $(\mathbf{P}_{\mathbf{H}^{\mathrm{der}}}, \mathbf{S}_{\mathbf{H}^{\mathrm{der}}}, K_{\mathbf{H}^{\mathrm{der}}})$ is a Siegel triple for $\mathbf{H}^{\mathrm{der}}$.

The unipotent radical of $\mathbf{P}_{\mathbf{H}^{\mathrm{der}}}$ is $\mathbf{U}_{\mathbf{H}^{\mathrm{der}}} = \mathbf{U}_1.\mathbf{U}_2$. The isogeny $\mathbf{H}_1 \times \mathbf{H}_2 \to \mathbf{H}^{\mathrm{der}}$ induces central isogenies $Z_{\mathbf{H}_1}(\mathbf{S}_1) \times Z_{\mathbf{H}_2}(\mathbf{S}_2) \to Z_{\mathbf{H}^{\mathrm{der}}}(\mathbf{S}_{\mathbf{H}^{\mathrm{der}}})$ and $\mathbf{P}_1/\mathbf{U}_1 \times \mathbf{P}_2/\mathbf{U}_2 \to \mathbf{P}_{\mathbf{H}^{\mathrm{der}}}/\mathbf{U}_{\mathbf{H}^{\mathrm{der}}}$, with the latter being defined over $\mathbb{Q}$. Hence the maximal $\mathbb{Q}$-anisotropic subgroup of $\mathbf{P}_{\mathbf{H}^{\mathrm{der}}}/\mathbf{U}_{\mathbf{H}^{\mathrm{der}}}$ is the product of the corresponding subgroups in $\mathbf{P}_1/\mathbf{U}_1$ and $\mathbf{P}_2/\mathbf{U}_2$. Lifting to $Z_{\mathbf{H}^{\mathrm{der}}}(\mathbf{S}_{\mathbf{H}^{\mathrm{der}}})$, it follows that $\mathbf{M}_{\mathbf{H}^{\mathrm{der}}} = \mathbf{M}_1.\mathbf{M}_2$.

The set of simple roots of $\mathbf{H}^{\mathrm{der}}$ (with respect to $(\mathbf{P}_{\mathbf{H}^{\mathrm{der}}}, \mathbf{S}_{\mathbf{H}^{\mathrm{der}}})$) is the union of the sets of simple roots of $\mathbf{H}_1$ and $\mathbf{H}_2$. It follows that $A_{\mathbf{H}^{\mathrm{der}},t} = A_{1,t}.A_{2,t}$.

We have $\mathfrak{S}_i = \Omega_i.A_{i,t}.K_i$ for $i = 1$ or 2, where Ω_i is a compact subset of $\mathbf{U}_i(\mathbb{R}).\mathbf{M}_i(\mathbb{R})^+$. Since $\mathbf{M}_1$ commutes with $\mathbf{U}_2$, $\Omega_1.\Omega_2$ is a compact subset of $\mathbf{U}_{\mathbf{H}^{\mathrm{der}}}(\mathbb{R}).\mathbf{M}_{\mathbf{H}^{\mathrm{der}}}(\mathbb{R})^+$. Hence

$$\mathfrak{S}_{\mathbf{H}^{\mathrm{der}}} = \Omega_1.\Omega_2.A_{1,t}.A_{2,t}.K_1.K_2$$

is a Siegel set in $\mathbf{H}^{\mathrm{der}}(\mathbb{R})$. Because $\mathbf{H}_1$ commutes with $\mathbf{H}_2$, we conclude that $\mathfrak{S}_{\mathbf{H}^{\mathrm{der}}} = \mathfrak{S}_1.\mathfrak{S}_2$. $\qquad\square$

Proof of Lemma 2.3. Write $\mathcal{F}_i = C_i.\mathfrak{S}_i^+.x_i$ as in the definition of Siegel fundamental sets.

By Lemma 2.4, $\mathfrak{S}_1.\mathfrak{S}_2$ is a Siegel set in $\mathbf{H}^{\mathrm{der}}(\mathbb{R})$. By [Orr18, Theorem 4.1], there exists a Siegel set $\mathfrak{S} \subset \mathbf{G}(\mathbb{R})$ and a finite set $C \subset \mathbf{G}(\mathbb{Q})$ such that $\mathfrak{S}_1.\mathfrak{S}_2 \subset C.\mathfrak{S}$.

Let $C_+ = C \cap \mathbf{G}(\mathbb{Q})_+$ and $\mathfrak{S}^+ = \mathfrak{S} \cap \mathbf{G}(\mathbb{R})^+$. Enlarging C and $\mathfrak{S}$ if necessary, we may ensure that $C_+.\mathfrak{S}^+$ is a fundamental set for Γ in $\mathbf{G}(\mathbb{R})^+$.

Let x_0 be the image of (x_1, x_2) under the inclusion $X_1^+ \times X_2^+ \to X^+$. Let $K_i = \mathrm{Stab}_{\mathbf{H}_i(\mathbb{R})}(x_i)$. Examining the proof of [Orr18, Theorem 4.1], we see that the maximal compact subgroup $K_{\mathbf{G}}$ in the Siegel triple used to construct $\mathfrak{S}$ can be chosen to be any maximal compact subgroup of $\mathbf{G}(\mathbb{R})$ which contains $K_1.K_2$. In particular, we can choose $K_{\mathbf{G}}$ to stabilise x_0.

It follows that $\mathcal{F} = C_+.\mathfrak{S}^+.x_0$ is a fundamental set for Γ in X^+ which contains $\mathfrak{S}_1^+.x_1 \times \mathfrak{S}_2^+.x_2$. Replacing C by $C_1.C_2.C$, we get $\mathcal{F}_1 \times \mathcal{F}_2 \subset \mathcal{F}$. $\qquad\square$

2.H. **Heights and determinants.** Let $k \geq 1$ be an integer. For any real number y, we define its k-**height** as

$$\mathrm{H}_k(y) := \min\{\max\{|a_0|, ..., |a_k|\} : a_i \in \mathbb{Z}, \ \gcd\{a_0, ..., a_k\} = 1, \ a_k y^k + ... + a_0 = 0\},$$

where we use the convention that, if the set is empty, that is, y is not algebraic of degree less than or equal to k, then $\mathrm{H}_k(y)$ is $+\infty$. For $y = (y_1, ..., y_m) \in \mathbb{R}^m$, we set

$$\mathrm{H}_k(y) := \max\{\mathrm{H}_k(y_1), ..., \mathrm{H}_k(y_m)\}.$$

We extend this definition to $\mathbb{C}^m$ by identifying it with $\mathbb{R}^{2m}$, taking real and imaginary parts. The 1-height of a matrix $g \in \mathrm{M}_n(\mathbb{Q})$ is the height of g considered as an element of $\mathbb{Q}^{n^2}$. For any matrix $g \in \mathrm{M}_n(\mathbb{Q})$, we write

$$\mathrm{det}^*(g) = |\det(g)| \cdot (\mathrm{lcm}\{b_{ij} : 1 \leq i, j \leq n\})^n \in \mathbb{Z},$$

where the entries of g written in lowest terms are a_{ij}/b_{ij}.

3. Complexities

We define a notion of complexity $\Delta(Z)$ for special subvarieties in a Hecke–facteur family which is similar to the general definition of complexity of special subvarieties from [DR18], but modified to be more convenient for the case of Hecke–facteur families. We will then define a second notion of complexity $\Delta'(Z)$ (which is even more specialised to the $E \times \mathrm{CM}$ case) and show that Δ and Δ' are polynomially bounded in terms of each other.

Let $(\mathbf{G}, X^+)$ be a Shimura datum component, let $K \subset \mathbf{G}(\mathbb{A}_f)$ be a compact open subgroup and let S be the Shimura variety component $\mathrm{Sh}_K(\mathbf{G}, X^+)$. Choose a faithful representation $\mathbf{G} \to \mathbf{GL}_{m,\mathbb{Q}}$ (we shall use this representation to talk about the height or det^* of elements of $\mathbf{G}(\mathbb{Q})$).

For each special point $s \in S$, one can define the following objects and quantities associated with s (as in [DR18, Definition 10.1]):

(1) $\mathbf{T} \subset \mathbf{G}$ is the Mumford–Tate group of s.
(2) $K_{\mathbf{T}}^m$ is the maximal compact open subgroup of $\mathbf{T}(\mathbb{A}_f)$. (There is a unique maximal compact open subgroup because $\mathbf{T}$ is a torus.)
(3) $D_{\mathbf{T}}$ is the absolute value of the discriminant of the splitting field of $\mathbf{T}$.

(4) $\Delta(s) = \max\{D_{\mathbf{T}}, \; [K_{\mathbf{T}}^m : K \cap \mathbf{T}(\mathbb{A}_f)]\}$.

The $\mathbb{Q}$-torus $\mathbf{T}$ is defined up to conjugation by $K \cap \mathbf{G}(\mathbb{Q})_+$, and $D_{\mathbf{T}}$ and $\Delta(s)$ are independent of the choice of $\mathbf{T}$ in its conjugacy class.

Let $(\mathbf{H}, X_{\mathbf{H}}^+, \mathbf{H}_1, \mathbf{H}_2)$ be a facteur datum for $(\mathbf{G}, X^+)$. For any special subvariety $Z \subset S$ in the associated Hecke–facteur family, we define the following quantities:

(1) $N(Z)$ is the smallest positive integer N such that there exist $\gamma \in \mathbf{G}(\mathbb{Q})_+$ and $x_2 \in X_2^+$ with $Z = Z_{\gamma,x_2}$ and $\det^*(\gamma) = N$.

(2) $\Delta(Z) = \max\{N(Z), \min\{\Delta(s) : s \in Z \text{ is a special point}\}\}$.

We call $\Delta(Z)$ the **complexity** of Z.

In the examples from section 2.E, the following lemma shows that the quantity $N(Z)$ can be interpreted as the smallest positive integer N such that, for every point $s \in Z$, there exist abelian varieties A_1 and A_2 (of dimensions d and e, respectively) and a polarised isogeny $A_1 \times A_2 \to A_s$ of degree N.

Lemma 3.1. *Let d, e and g be positive integers such that $d + e = g$ and let $(\mathbf{H}, X_{\mathbf{H}}^+, \mathbf{H}_1, \mathbf{H}_2)$ be the facteur datum described in section 2.E. Let Z be a special subvariety of $\mathcal{A}_g$ in the Hecke–facteur family associated with $(\mathbf{H}, X_{\mathbf{H}}^+, \mathbf{H}_1, \mathbf{H}_2)$.*

For each positive integer N, the following statements are equivalent:

(i) There exists $\gamma \in \mathbf{GSp}_{2g}(\mathbb{Q})_+$ and a pre-special point $x_2 \in X_2^+$ such that $Z = Z_{\gamma,x_2}$ and $\det^*(\gamma) = N$ (with respect to the inclusion $\mathbf{GSp}_{2g} \hookrightarrow \mathbf{GL}_{2g}$).

(ii) There exists a pre-special point $x_2 \in X_2^+$ such that, for every point $s \in Z$, there exists an abelian variety A_1 of dimension d and a polarised isogeny $A_1 \times A_{x_2} \to A_s$ of degree N, where A_{x_2} denotes the abelian variety of dimension e naturally associated with $x_2 \in X_2^+ = \mathcal{H}_e$.

(iii) There exists a pre-special point $x_2 \in X_2^+$, a point $s \in Z$ such that Z is the smallest special subvariety containing s, an abelian variety A_1 of dimension d and a polarised isogeny $A_1 \times A_{x_2} \to A_s$ of degree N.

Proof. First suppose that we are given γ and x_2 as in (i). We can multiply γ by the lowest common multiple of the denominators of its entries without changing $\det^*(\gamma)$, so we may assume that $\gamma \in \mathbf{GSp}_{2g}(\mathbb{Q})_+ \cap \mathrm{M}_{2g}(\mathbb{Z})$. By the definition of Z_{γ,x_2}, for each point $s \in Z$, we have $s = \pi(\gamma.(x_1, x_2))$ for some point $x_1 \in X_1^+$. Then the matrix γ is the rational representation of a polarised isogeny $A_{x_1} \times A_{x_2} \to A_s$ of degree $\det(\gamma) = N$, where A_{x_1} denotes the abelian variety of dimension d naturally associated with $x_1 \in X_1^+ = \mathcal{H}_d$. This yields (ii).

It is obvious that (ii) implies (iii).

Finally, suppose that (iii) holds. Let p_1, p_2 denote the first and second projections $\mathcal{H}_g \times \mathcal{H}_g \to \mathcal{H}_g$ and let π denote the uniformising map $\mathcal{H}_g \to \mathcal{A}_g$. Let

$$T_N = \{(s_1, s_2) \in \mathcal{A}_g \times \mathcal{A}_g : \exists \text{ a polarised isogeny } A_{s_1} \to A_{s_2} \text{ of degree } N\}$$

and let $X_N = (\pi \times \pi)^{-1}(T_N)$. Let

$$W_N = \Big((X_1^+ \times \{x_2\}) \times \mathcal{H}_g\Big) \cap X_N.$$

Each irreducible component of X_N is of the form $X_\gamma = \{(x, \gamma x) : x \in \mathcal{H}_g\}$ for some $\gamma \in \mathbf{GSp}_{2g}(\mathbb{Q})_+ \cap \mathrm{M}_{2g}(\mathbb{Z})$ such that $\det(\gamma) = N$. Hence, each irreducible component of W_N has the form

$$W_\gamma = \{((x_1, x_2), \gamma(x_1, x_2)) : x_1 \in \mathcal{H}_d\}$$

for some γ of the same type.

By (iii), $s \in (\pi \circ p_2)(W_N)$. Each irreducible component of $(\pi \circ p_2)(W_N)$ is a special subvariety of $\mathcal{A}_g$. Since Z is the smallest special subvariety containing s, we deduce that $Z \subset (\pi \circ p_2)(W_N)$. Since Z is irreducible, it is contained in $(\pi \circ p_2)(W_\gamma)$ for some γ. Because Z is in the Hecke–facteur family, $\dim(Z) = \dim(W_\gamma)$. Therefore, $Z = (\pi \circ p_2)(W_\gamma) = Z_{\gamma, x_2}$. $\qquad\square$

Now suppose that $(\mathbf{H}, X_{\mathbf{H}}^+, \mathbf{H}_1, \mathbf{H}_2)$ is the facteur datum discussed in section 2.E with $g = 2$ and $d = e = 1$, that is, the datum associated with $E \times$ CM curves. In this case, we can relate the above definition of complexity to a simpler definition, which is crucial for the proof of Theorem 1.4. The new definition of complexity relies on the following lemma.

Lemma 3.2. *Let (A, λ) be a principally polarised abelian surface which is isogenous to a product of elliptic curves $E_1 \times E_2$ where E_1 and E_2 are not isogenous to each other. Let N be the smallest positive integer such that there exist elliptic curves E_1 and E_2 and an isogeny $E_1 \times E_2 \to A$ of degree N.*

Then the pair of elliptic curves (E_1, E_2) such that there exists an isogeny $E_1 \times E_2 \to A$ of degree N is unique up to swapping E_1 and E_2.

Proof. Let E_1 and E_2 be elliptic curves for which there exists an isogeny $\varphi \colon E_1 \times E_2 \to A$ of degree N. Let $\Lambda = H_1(A, \mathbb{Z})$ and $\Lambda_i = H_1(E_i, \mathbb{Z})$. Let ψ be the symplectic form on Λ induced by λ and let ψ_i be the symplectic form on Λ_i induced by the unique principal polarisation of E_i.

Because $\deg \varphi$ is as small as possible, we must have $\varphi' = \varphi$ in Lemma 2.1. Looking at the proof of Lemma 2.1, we see that there is a positive integer n such that $\psi_{|\Lambda_i} = n\psi_i$. Then $N = \deg \varphi = n^2$.

As a first step, we show that $\ker(\varphi) \subset (E_1 \times E_2)[n]$.

Choose a basis $\{x, y\}$ for Λ_1 such that $\psi_1(x, y) = 1$. For any $v \in \Lambda$, we can write $v = v_1 + v_2$ where $v_i \in \Lambda_i \otimes \mathbb{Q}$. We can calculate

$$\psi(v, x) = \psi(v_1, x) = n\psi_1(v_1, x).$$

Since $\psi(v, x) \in \mathbb{Z}$, $\psi_1(v_1, x) \in \frac{1}{n}\mathbb{Z}$. Similarly $\psi_1(v_1, y) \in \frac{1}{n}\mathbb{Z}$. Now

$$v_1 = \psi_1(v_1, y).x - \psi_1(v_1, x).y$$

and so $v_1 \in \frac{1}{n}\Lambda_1$.

A similar argument shows that $v_2 \in \frac{1}{n}\Lambda_2$. Thus $\Lambda \subset \frac{1}{n}\Lambda_1 + \frac{1}{n}\Lambda_2$. In other words, n annihilates $\Lambda/(\Lambda_1 + \Lambda_2) \cong \ker(\varphi)$.

Suppose that there exists another pair of elliptic curves E_1', E_2' and an isogeny $\varphi' \colon E_1' \times E_2' \to A$ also of degree N.

We can apply the above argument to φ', deducing that $\ker(\varphi') \subset (E_1' \times E_2')[n]$. Consequently, there exists an isogeny $\psi' \colon A \to E_1' \times E_2'$ such that $\psi' \circ \varphi' = [n]_{E_1' \times E_2'}$. We will consider the isogeny $\psi' \circ \varphi \colon E_1 \times E_2 \to E_1' \times E_2'$.

Label E_1' and E_2' so that they are isogenous to E_1 and E_2 in that order. Because E_1 is not isogenous to E_2, we have $\operatorname{Hom}(E_1, E_2') = \operatorname{Hom}(E_2, E_1') = 0$. Consequently

$$\psi' \circ \varphi = (\alpha_1, \alpha_2)$$

for some isogenies $\alpha_1 \colon E_1 \to E_1'$ and $\alpha_2 \colon E_2 \to E_2'$.

Let $G = \ker(\varphi) \subset (E_1 \times E_2)[n]$ and let p_i denote the projection $(E_1 \times E_2)[n] \to E_i[n]$ for $i = 1$ and 2.

Observe that $\varphi \colon E_1 \times E_2 \to A$ factors through $(E_1 \times E_2)/(G \cap E_2[n])$ because $G \cap E_2[n] \subset G = \ker(\varphi)$. Hence the minimality of $\deg \varphi$ implies that $G \cap E_2[n] = 0$. We conclude that $p_{1|G}$ is injective. It follows that $\# p_1(G) = \# G = n^2$. Thus $p_1(G) = E_1[n]$. Similarly $p_2(G) = E_2[n]$.

If $g = (g_1, g_2) \in G$, then $(\alpha_1(g_1), \alpha_2(g_2)) = \psi' \varphi(g) = 0$ and so $p_i(G) \subset \ker(\alpha_i)$.

Since $\deg(\varphi') = n^2$ and $\deg[n]_{E_1' \times E_2'} = n^4$, we have $\deg(\psi') = n^2$. Consequently $\deg(\psi' \circ \varphi) = n^4$. Meanwhile $\ker(\psi' \circ \varphi) = \ker(\alpha_1) \times \ker(\alpha_2)$ so the inclusions $p_i(G) \subset \ker(\alpha_i)$ must be equalities.

It follows that $E_i' \cong E_i/E_i[n] \cong E_i$. $\square$

We continue to consider the $E \times$ CM Hecke–facteur family, using the standard representation $\mathbf{GSp}_4 \to \mathbf{GL}_4$ to define $\det^*$. Let Z be a special curve in this family. By Lemma 3.1, there exists $x_2 \in X_2^+$ such that, for every $s \in Z$, there exists an elliptic curve $E_1(s)$ and an isogeny $E_1(s) \times E_{x_2} \to A_s$ of degree $N(Z)$, where E_{x_2} is the elliptic curve naturally associated with $x_2 \in X_2^+ = \mathcal{H}_1$. Furthermore, for any point $s \in Z \cap \Sigma$, $N(Z)$ is the smallest degree of an isogeny of A_s to a product of elliptic curves. Therefore, by Lemma 3.2, $E_1(s)$ and E_{x_2} are the unique elliptic curves whose product has an isogeny to A_s of degree $N(Z)$. Since E_{x_2} has CM while $E_1(s)$, E_{x_2} is uniquely associated with Z. Hence, it makes sense to define

$$\Delta'(Z) = \max\{N(Z), |\operatorname{disc}(\operatorname{End}(E_{x_2}))|\}.$$

Remark 3.3. If s is an $E \times$ CM point in the base of a principally polarised abelian scheme $\mathfrak{A} \to V$, as defined in section 1.B, then the point of $\mathcal{A}_2$ attached to the fiber $\mathfrak{A}_s$ lies in a unique $E \times$ CM curve Z. In this situation (as for example in the statement of Conjecture 1.5), we write $N(s) = N(Z)$ and $\Delta'(s) = \Delta'(Z)$.

For a special curve Z in the $E \times$ CM Hecke–facteur family, our two notions of complexity $\Delta(Z)$ and $\Delta'(Z)$ are polynomially bounded with respect to each other. Before proving this, we shall prove two lemmas. The first is a special case of the reverse of [Tsi12, Lemma 7.2] or [DO16, Theorem 4.1]. (We expect that the reverse of [DO16, Theorem 4.1] should hold for all special points in an arbitrary Shimura variety, but the proof is likely to be somewhat more complicated.)

Lemma 3.4. *There exist absolute constants C_7 and C_8 such that for every special point $s \in \mathcal{A}_2$, if there exists a polarised isogeny $E_1 \times E_2 \to A_s$ where E_1 and E_2 are non-isogenous elliptic curves both of which have CM, then*

$$\Delta(s) \leq C_7 \left| \mathrm{disc}(\mathrm{End}(A_s)) \right|^{C_8}.$$

Proof. Let $F_i = \mathrm{End}(E_i) \otimes \mathbb{Q}$ for $i = 1$ and 2. Let $\mathcal{O}_i$ denote the ring of integers of F_i and let $\mathcal{O} = \mathcal{O}_1 \times \mathcal{O}_2$, $\mathcal{O}_p = \mathcal{O} \otimes \mathbb{Z}_p$ and $\hat{\mathcal{O}} = \prod_p \mathcal{O}_p$. Let $R = \mathrm{End}(A_s)$, $R_p = R \otimes \mathbb{Z}_p$ and $\hat{R} = \prod_p R_p$. Observe that R is an order in $F_1 \times F_2$ so

$$\mathrm{disc}(R) = \mathrm{disc}(F_1) \, \mathrm{disc}(F_2) [\mathcal{O} : R]^2.$$

We use the notation $\mathbf{T}$, $K_{\mathbf{T}}^m$ and $D_{\mathbf{T}}$ as in the definition of $\Delta(s)$ in section 3.

Fix a basis for $H_1(A_s, \mathbb{Z})$. With respect to this basis, we get a homomorphism of rings $\iota \colon R \to \mathrm{M}_4(\mathbb{Z})$ and an injection $\mathbf{T} \to \mathbf{GL}_{4,\mathbb{Q}}$.

Because the Mumford–Tate group of an abelian surface is always as large as possible given its endomorphism ring and polarisation (see section 2.F), $\mathbf{T}$ is equal to the intersection of $\mathbf{GSp}_4$ with the centraliser (in $\mathbf{GL}_4$) of $\iota(R)$. Now $\iota(R \otimes \mathbb{Q})$ is a commutative algebra of dimension 4, so it is its own centraliser in $\mathrm{M}_4(\mathbb{Q})$. Hence

$$\mathbf{T}(\mathbb{Q}) = \iota((R \otimes \mathbb{Q})^\times) \cap \mathbf{GSp}_4(\mathbb{Q}).$$

We deduce that $\mathbf{T}$ is a subtorus of $\mathrm{Res}_{F_1/\mathbb{Q}} \mathbb{G}_m \times \mathrm{Res}_{F_2/\mathbb{Q}} \mathbb{G}_m$ and hence is split over the compositum $F_1 F_2$. Consequently

$$D_{\mathbf{T}} \leq |\mathrm{disc}(F_1 F_2)| \leq |\mathrm{disc}(F_1)|^2 |\mathrm{disc}(F_2)|^2 \leq |\mathrm{disc}(R)|^2$$

(the middle inequality is well-known; see for example [Jar14, Exercise 8.10]).

Let

$$K_{\mathbf{T}} = \mathbf{T}(\mathbb{A}_f) \cap \mathbf{GSp}_4(\hat{\mathbb{Z}}) = \mathbf{T}(\mathbb{A}_f) \cap \iota(\hat{R}^\times).$$

The maximal compact subgroup of $((F_1 \times F_2) \otimes \mathbb{A}_f)^\times$ is $\hat{\mathcal{O}}^\times$, so $K_{\mathbf{T}}^m = \mathbf{T}(\mathbb{A}_f) \cap \iota(\hat{\mathcal{O}}^\times)$. Therefore

$$[K_{\mathbf{T}}^m : K_{\mathbf{T}}] = [\mathbf{T}(\mathbb{A}_f) \cap \iota(\hat{\mathcal{O}}^\times) : \mathbf{T}(\mathbb{A}_f) \cap \iota(\hat{R}^\times)] \leq [\hat{\mathcal{O}}^\times : \hat{R}^\times].$$

Thus it will suffice to show that

$$[\hat{\mathcal{O}}^\times : \hat{R}^\times] \leq [\mathcal{O} : R]^4.$$

We will prove this prime by prime: for each prime p we will show that

$$[\mathcal{O}_p^\times : R_p^\times] \leq [\mathcal{O}_p : R_p]^4. \tag{1}$$

Let $N_p = [\mathcal{O}_p : R_p]$. If $\mathcal{O}_p = R_p$, then (1) is obvious. So we may assume that $\mathcal{O}_p \neq R_p$ and then p divides N_p.

Since $1 \in R_p$ and $N_p \mathcal{O}_p \subset R_p$, we have $1 + N_p \mathcal{O}_p \subset R_p$. Because p divides N_p, if $x \in 1 + N_p \mathcal{O}_p$ then $x \in \mathcal{O}_p^\times$ and x^{-1} is also congruent to 1 mod N_p. In other words $x^{-1} \in 1 + N_p \mathcal{O}_p \subset R_p$ and so $x \in R_p^\times$. This proves that $1 + N_p \mathcal{O}_p \subset R_p^\times$. Hence

$$[\mathcal{O}_p^\times : R_p^\times] \leq [\mathcal{O}_p^\times : 1 + N_p \mathcal{O}_p].$$

Finally the inclusion $\mathcal{O}_p^\times \to \mathcal{O}_p$ induces an injective map of sets $\mathcal{O}_p^\times/(1+N_p\mathcal{O}_p) \to \mathcal{O}_p/N_p\mathcal{O}_p$ and so

$$[\mathcal{O}_p^\times : 1 + N_p\mathcal{O}_p] \leq [\mathcal{O}_p : N_p\mathcal{O}_p] = N_p^4.$$

This completes the proof. $\square$

Our second lemma is a general result on endomorphisms and isogenies of abelian varieties. Write $Z(R)$ for the centre of a ring R. (Note that the analogous statement to Lemma 3.5 still holds if you replace $Z(\mathrm{End}(A))$ and $Z(\mathrm{End}(B))$ by $\mathrm{End}(A)$ and $\mathrm{End}(B)$ respectively, with the same modification to the proof.)

Lemma 3.5. *Let A and B be abelian varieties such that that there exists an isogeny $\varphi \colon A \to B$ of degree N, and let $r = \mathrm{rk}_\mathbb{Z}(Z(\mathrm{End}(A)))$. Then*

$$|\mathrm{disc}(Z(\mathrm{End}(B)))| \leq N^{2r}|\mathrm{disc}(Z(\mathrm{End}(A)))|.$$

Proof. Equip $Z(\mathrm{End}(A)) \otimes \mathbb{Q}$ and $Z(\mathrm{End}(B)) \otimes \mathbb{Q}$ with the symmetric bilinear forms $(x,y) \mapsto \mathrm{tr}(xy)$, where tr denotes the trace on the respective $\mathbb{Q}$-algebras. We denote these forms ψ_A and ψ_B. Recall that by definition, $\mathrm{disc}(Z(\mathrm{End}(A)))$ is the discriminant of the bilinear form ψ_A on $Z(\mathrm{End}(A))$, and similarly for B.

The isogeny φ induces an isomorphism $\iota \colon \mathrm{End}(A) \otimes \mathbb{Q} \to \mathrm{End}(B) \otimes \mathbb{Q}$ given by $\iota(\alpha) = \varphi\alpha\varphi^{-1}$. This restricts to an isomorphism $Z(\mathrm{End}(A)) \otimes \mathbb{Q} \to Z(\mathrm{End}(B)) \otimes \mathbb{Q}$. Any isomorphism of $\mathbb{Q}$-algebras preserves the trace, and therefore is compatible with the bilinear forms ψ_A and ψ_B.

Because there exists an isogeny $\psi \colon B \to A$ such that $\psi \circ \varphi = [N]$, ι maps $N \cdot \mathrm{End}(A)$ into $\mathrm{End}(B)$, and hence $N \cdot Z(\mathrm{End}(A))$ into $Z(\mathrm{End}(B))$. Consequently

$$|\mathrm{disc}(\psi_B|Z(\mathrm{End}(B)))| \leq |\mathrm{disc}(\psi_A|N \cdot Z(\mathrm{End}(A)))| = N^{2r}|\mathrm{disc}(\psi_A|Z(\mathrm{End}(A)))|.$$

$\square$

Now we combine the above two lemmas to compare Δ and Δ'.

Lemma 3.6. *There exist positive constants C_9, C_{10}, C_{11}, and C_{12} such that, if Z is a special curve in the $E \times$ CM Hecke-facteur family, then*

$$C_9\Delta(Z)^{C_{10}} \leq \Delta'(Z) \leq C_{11}\Delta(Z)^{C_{12}}.$$

Proof. Write $Z = Z_{\gamma,x_2}$ where $\gamma \in \mathbf{GSp}_4(\mathbb{Q})_+ \cap \mathrm{M}_4(\mathbb{Z})$, $x_2 \in X_2^+$ and $\det^*(\gamma) = N(Z)$.

Fix two CM elliptic curves E_1 and E_1' which are not isogenous to each other. Assume that E_{x_2} is not isogenous to E_1 (if E_{x_2} is isogenous to E_1, then use E_1' instead of E_1 and apply the same argument).

Let x_1 be a point in X_1^+ corresponding to the elliptic curve E_1, and let s denote the image of $\gamma(x_1, x_2)$ in $\mathcal{A}_2$. Then s is a special point contained in Z, so

$$\Delta(Z) \leq \max\{N(Z), \Delta(s)\}.$$

By Lemma 3.4, we have

$$\Delta(s) \leq C_7 \, |\mathrm{disc}(\mathrm{End}(A_s))|^{C_8}.$$

Now γ is the rational representation of a polarised isogeny $E_1 \times E_{x_2} \to A_s$ of degree $\det^*(\gamma) = N(Z)$. Noting that $\mathrm{rk}_{\mathbb{Z}}(\mathrm{End}(E_1 \times E_{x_2})) = 4$ and that $\mathrm{End}(E_1 \times E_{x_2})$ and $\mathrm{End}(A_s)$ are commutative, Lemma 3.5 tells us that

$$|\mathrm{disc}(\mathrm{End}(A_s))| \leq N(Z)^8 |\mathrm{disc}(\mathrm{End}(E_1 \times E_{x_2}))|.$$

Since E_{x_2} is not isogenous to E_1, $\mathrm{End}(E_1 \times E_{x_2}) = \mathrm{End}(E_1) \times \mathrm{End}(E_{x_2})$ and so

$$|\mathrm{disc}(\mathrm{End}(E_1 \times E_{x_2}))| = |\mathrm{disc}(\mathrm{End}(E_1)) \, \mathrm{disc}(\mathrm{End}(E_{x_2}))| \leq C_{13} \Delta'(Z)$$

where $C_{13} = \max\{|\mathrm{disc}(\mathrm{End}(E_1))|, |\mathrm{disc}(\mathrm{End}(E_1'))|\}$.

Combining the above inequalities gives

$$C_9 \Delta(Z)^{C_{10}} \leq \Delta'(Z).$$

To prove the second inequality of the lemma, let $s \in Z$ be a special point such that $\Delta(s)$ is minimal. By Lemma 3.1, there exists an elliptic curve $E_1(s)$ and an isogeny $E_1(s) \times E_{x_2} \to A_s$ of degree $N(Z)$. Since both $E_1(s) \times E_{x_2}$ and A_s are principally polarised, we deduce that there exists an isogeny in the reverse direction $A_s \to E_1(s) \times E_{x_2}$, also of degree $N(Z)$. Hence Lemma 3.5 implies that

$$|\mathrm{disc}(Z(\mathrm{End}(E_1(s) \times E_{x_2})))| \leq N(Z)^8 |\mathrm{disc}(Z(\mathrm{End}(A_s)))|.$$

Let $\mathcal{O}_s$ denote the maximal order in $Z(\mathrm{End}(A_s)) \otimes \mathbb{Q}$ (which is either an imaginary quadratic field or a product of two imaginary quadratic fields). Then

$$\mathrm{disc}(Z(\mathrm{End}(A_s))) = [\mathcal{O}_s : Z(\mathrm{End}(A_s))]^2 \, \mathrm{disc}(Z(\mathrm{End}(A_s)) \otimes \mathbb{Q}).$$

By [Tsi12, Lemma 7.2], we have (using the notation from the definition of $\Delta(s)$)

$$[\mathcal{O}_s : Z(\mathrm{End}(A_s))] \leq [K_{\mathbf{T}}^m : \mathbf{GSp}_{2g}(\hat{\mathbb{Z}}) \cap \mathbf{T}(\mathbb{A}_f)]^{C_{14}} D_{\mathbf{T}}^{C_{15}} \leq \Delta(s)^{C_{16}}.$$

Combining the above inequalities, we conclude that

$$|\mathrm{disc}(Z(\mathrm{End}(E_1(s) \times E_{x_2})))| \leq \Delta(s)^{C_{17}} |\mathrm{disc}(Z(\mathrm{End}(A_s)) \otimes \mathbb{Q})|. \qquad (2)$$

We now split into two cases depending on whether $E_1(s)$ is isogenous to E_{x_2} or not. If they are not isogenous, then $F_1(s) = \mathrm{End}(E_1(s)) \otimes \mathbb{Q}$ and $F_2 = \mathrm{End}(E_{x_2}) \otimes \mathbb{Q}$ are distinct imaginary quadratic fields and we have

$$Z(\mathrm{End}(E_1(s) \times E_{x_2})) = \mathrm{End}(E_1(s)) \times \mathrm{End}(E_{x_2})$$

so that

$$|\mathrm{disc}(\mathrm{End}(E_{x_2}))| \leq |\mathrm{disc}(Z(\mathrm{End}(E_1(s) \times E_{x_2})))|$$

while

$$|\mathrm{disc}(Z(\mathrm{End}(A_s)) \otimes \mathbb{Q})| = |\mathrm{disc}(F_1(s)) \, \mathrm{disc}(F_2)| \leq |\mathrm{disc}(F_1(s)F_2)|^2 = D_{\mathbf{T}}^2.$$

Combining these with (2) completes the proof of the second inequality in this case.

If $E_1(s)$ is isogenous to E_{x_2}, then $\mathrm{End}(E_1(s) \times E_{x_2}) \otimes \mathbb{Q} \cong \mathrm{M}_2(F)$ where $F = \mathrm{End}(E_1(s)) \otimes \mathbb{Q} \cong \mathrm{End}(E_{x_2}) \otimes \mathbb{Q}$. Hence $Z(\mathrm{End}(E_1(s) \times E_{x_2}))$ is an order in F. If

$\iota \colon E_{x_2} \to E_1(s) \times E_{x_2}$ and $\beta \colon E_1(s) \times E_{x_2} \to E_{x_2}$ denote the inclusion and projection morphisms, then $\alpha \mapsto \beta \alpha \iota$ is an injection $Z(\mathrm{End}(E_1(s) \times E_{x_2})) \to \mathrm{End}(E_{x_2})$. Hence

$$|\mathrm{disc}(\mathrm{End}(E_{x_2}))| \leq |\mathrm{disc}(Z(\mathrm{End}(E_1(s) \times E_{x_2})))|.$$

The splitting field of $\mathbf{T}$ is $F \cong Z(\mathrm{End}(A_s) \otimes \mathbb{Q}$, so

$$|\mathrm{disc}(Z(\mathrm{End}(A_s)) \otimes \mathbb{Q})| = D_{\mathbf{T}}.$$

Again combining these with (2) completes the proof. □

4. Galois action on Hecke–facteur special subvarieties

Let S be a Shimura variety component defined over a number field E_S. Then $\mathrm{Gal}(\overline{\mathbb{Q}}/E_S)$ acts on the set of $\overline{\mathbb{Q}}$-subvarieties of S. We shall show that, after restricting to $\mathrm{Gal}(\overline{\mathbb{Q}}/F)$ for a suitable field extension F/E_S, this action permutes the special subvarieties in a given Hecke–facteur family.

The main result of this section (Proposition 4.3) shares a certain similarity with [DR18, Conjecture 12.6]. Indeed, [DR18, Conjecture 12.6] predicts that, for any special subvariety Z in a suitable facteur family, we can make an extension of the base field (whose degree is very small relative to the complexity of Z) over which all Galois conjugates of Z are members of finitely many *facteur* families. Proposition 4.3 is weaker than this because it only asserts that the Galois conjugates are members of the same *Hecke–facteur* family. On the other hand, Proposition 4.3 only involves an extension of the base field which is independent of Z.

Given a Shimura datum component $(\mathbf{G}, X^+)$, let X denote the $\mathbf{G}(\mathbb{R})$-conjugacy class of morphisms $\mathbb{S} \to \mathbf{G}(\mathbb{R})$ containing X^+. Then $(\mathbf{G}, X)$ is a Shimura datum. For any subset A of X and $a \in \mathbf{G}(\mathbb{A}_f)$, we will denote by $[A, a]_K$ the image of $A \times \{a\}$ in $\mathrm{Sh}_K(\mathbf{G}, X)(\mathbb{C})$, in analogy with a standard notation for points. In particular, the geometrically irreducible components of $\mathrm{Sh}_K(\mathbf{G}, X)$ are the subsets of the form $[X^+, a]_K$ for any $a \in \mathbf{G}(\mathbb{A}_f)$.

We begin with a lemma on the restriction of Hecke correspondences on $\mathrm{Sh}_K(\mathbf{G}, X)$ to the connected component $[X^+, 1]_K$. Recall that T_g denotes the map on algebraic cycles induced by the Hecke correspondence associated with g (see section 2.C).

Lemma 4.1. *Let $(\mathbf{G}, X^+)$ be a Shimura datum component and let $K \subset \mathbf{G}(\mathbb{A}_f)$ be a compact open subgroup. Let S denote the Shimura variety component $[X^+, 1]_K \subset \mathrm{Sh}_K(\mathbf{G}, X)$.*

Let Z be an irreducible complex algebraic subvariety of S and let Y be an irreducible component of $\pi^{-1}(Z)$. Let $g \in \mathbf{G}(\mathbb{Q})_+$. Then every irreducible component of $T_{g^{-1}}(Z) \cap S$ can be written in the form $[g'Y, 1]_K$ for some $g' \in \mathbf{G}(\mathbb{Q})_+$.

Proof. Each irreducible component of $T_{g^{-1}}(Z)$ has the form

$$Z' = [Y, kg^{-1}]_K$$

for some $k \in K$. If $Z' \subset S$, then $kg^{-1} \in \mathbf{G}(\mathbb{Q})_+ K$, or in other words

$$kg^{-1}k' \in \mathbf{G}(\mathbb{Q})_+$$

for some $k' \in K$. Letting $g' = (kg^{-1}k')^{-1}$, we get

$$Z' = [Y, g'^{-1}]_K = [g'Y, 1]_K. \qquad \square$$

Remark 4.2. Fix a faithful representation $\mathbf{G} \to \mathbf{GL}_{m,\mathbb{Q}}$ such that the image of K is contained in $\mathbf{GL}_m(\hat{\mathbb{Z}})$. Since $g' = k'^{-1}gk^{-1}$, we have $|\det(g')| = |\det(g)|$. We also claim that the lowest common multiples of the denominators of g' and g are equal. This is because g' is obtained from g by multiplying on the left and right by elements of K. Working prime by prime, we see that such operations cannot increase the lowest common multiple of the denominators. Since also $g = k'g'k$, these operations cannot decrease the lowest common multiple of the denominators either. We conclude that $\det^*(g') = \det^*(g)$.

Proposition 4.3. *Let* $(\mathbf{G}, X^+)$ *be a Shimura datum component and let* $K \subset \mathbf{G}(\mathbb{A}_f)$ *be a compact open subgroup. Let* S *denote the Shimura variety component* $[X^+, 1]_K \subset \mathrm{Sh}_K(\mathbf{G}, X)$.

Let $(\mathbf{H}, X_{\mathbf{H}}^+, \mathbf{H}_1, \mathbf{H}_2)$ *be a facteur datum for* $(\mathbf{G}, X^+)$ *and let* $\mathfrak{F}$ *denote the associated Hecke–facteur family of special subvarieties of* S.

Let E_S *denote the extension of* $E_{\mathbf{G}} := E(\mathbf{G}, X)$ *over which* S *is defined. Let* F *denote the compositum of* $E_{\mathbf{H}} := E(\mathbf{H}, X_{\mathbf{H}})$ *and* E_S.

For every special subvariety $Z \in \mathfrak{F}$, *and every* $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/F)$, *the Galois conjugate subvariety* $\sigma(Z)$ *is again a member of the Hecke–facteur family* $\mathfrak{F}$. *In other words,* $\sigma(Z) = Z_{g_\sigma, x_\sigma}$ *for some* $g_\sigma \in \mathbf{G}(\mathbb{Q})_+$ *and some pre-special point* $x_\sigma \in X_2^+$.

Proof. First, let $g \in \mathbf{G}(\mathbb{Q})_+$ and $x_2 \in X_2^+$ such that $Z = Z_{g,x_2}$. Now let $Z_{x_2} := [X_1^+, x_2]_{K_{\mathbf{H}}} \subseteq \mathrm{Sh}_{K_{\mathbf{H}}}(\mathbf{H}, X_{\mathbf{H}})$, where $K_{\mathbf{H}} := \mathbf{H}(\mathbb{A}_f) \cap K$, and let

$$\varphi_{\mathbf{H}} : \mathrm{Sh}_{K_{\mathbf{H}}}(\mathbf{H}, X_{\mathbf{H}}) \to \mathrm{Sh}_K(\mathbf{G}, X)$$

denote the natural morphism of Shimura varieties. By definition, Z_{g,x_2} is an irreducible component of $T_{g^{-1}}(\varphi_{\mathbf{H}}(Z_{x_2}))$. Since $T_{g^{-1}}$ is defined over $E_{\mathbf{G}} \subseteq F$ and $\varphi_{\mathbf{H}}$ is defined over $E_{\mathbf{H}} \subseteq F$, we see that, for any $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/F)$, the subvariety $\sigma(Z_{g,x_2})$ is an irreducible component of $T_{g^{-1}}(\varphi_{\mathbf{H}}(\sigma(Z_{x_2})))$.

Let $K_{\mathbf{H}}^{\mathrm{ad}} \subset \mathbf{H}^{\mathrm{ad}}(\mathbb{A}_f)$ denote a maximal compact open subgroup containing the image of $K_{\mathbf{H}}$ under the natural morphism $q : \mathbf{H}(\mathbb{A}_f) \to \mathbf{H}^{\mathrm{ad}}(\mathbb{A}_f)$. Then $K_{\mathbf{H}}^{\mathrm{ad}} = K_1 \times K_2$ for maximal compact subgroups $K_1 \subset \mathbf{H}_1^{\mathrm{ad}}(\mathbb{A}_f)$ and $K_2 \subset \mathbf{H}_2^{\mathrm{ad}}(\mathbb{A}_f)$ and we obtain a finite morphism of Shimura varieties

$$\varphi^{\mathrm{ad}} : \mathrm{Sh}_{K_{\mathbf{H}}}(\mathbf{H}, X_{\mathbf{H}}) \to \mathrm{Sh}_{K_{\mathbf{H}}^{\mathrm{ad}}}(\mathbf{H}^{\mathrm{ad}}, X_{\mathbf{H}}^{\mathrm{ad}}) = \mathrm{Sh}_{K_1}(\mathbf{H}_1^{\mathrm{ad}}, X_1) \times \mathrm{Sh}_{K_2}(\mathbf{H}_2^{\mathrm{ad}}, X_2),$$

which is defined over $E_{\mathbf{H}}$. We conclude that, for any $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/F)$,

$$\varphi^{\mathrm{ad}}(\sigma(Z_{x_2})) = \sigma(\varphi^{\mathrm{ad}}(Z_{x_2})) = \sigma([X_1^+, 1]_{K_1}) \times \sigma([\{x_2\}, 1]_{K_2}).$$

The first factor is a geometrically connected component of $\mathrm{Sh}_{K_1}(\mathbf{H}_1^{\mathrm{ad}}, X_1)$ because $\mathrm{Gal}(\overline{\mathbb{Q}}/F)$ acts on the set of such components. The second factor is a special point

of $\operatorname{Sh}_{K_2}(\mathbf{H}_2^{\mathrm{ad}}, X_2)$ by [Orr19, Theorem 5.1]. Hence

$$\varphi^{\mathrm{ad}}(\sigma(Z_{x_2})) = \sigma(\varphi^{\mathrm{ad}}(Z_{x_2})) = [X_1^+, a_\sigma]_{K_1} \times [\{y_\sigma\}, b_\sigma]_{K_2},$$

for some pre-special point $y_\sigma \in X_2^+$ and some $a_\sigma \in \mathbf{H}_1^{\mathrm{ad}}(\mathbb{A}_f)$ and $b_\sigma \in \mathbf{H}_2^{\mathrm{ad}}(\mathbb{A}_f)$. In other words, $\sigma(Z_{x_2})$ is an irreducible component of

$$(\varphi^{\mathrm{ad}})^{-1}([X_1^+, a_\sigma]_{K_1} \times [\{y_\sigma\}, b_\sigma]_{K_2}),$$

and so $\sigma(Z_{x_2})$ is equal to $[X_1^+ \times \{x_\sigma\}, h_\sigma]_{K_\mathbf{H}}$ for some pre-special point $x_\sigma \in X_2^+$ and some element $h_\sigma \in \mathbf{H}(\mathbb{A}_f)$ such that $q(h_\sigma) = (a_\sigma, b_\sigma)$.

On the other hand, because $E_S \subseteq F$, we have

$$\varphi_\mathbf{H}(\sigma(Z_{x_2})) = \sigma(\varphi_\mathbf{H}(Z_{x_2})) \subseteq S,$$

for any $\sigma \in \operatorname{Gal}(\overline{\mathbb{Q}}/F)$. Hence $[X^+, h_\sigma]_K = S$ and so $h_\sigma = h_{\sigma,\mathbb{Q}}k$ for some $h_{\sigma,\mathbb{Q}} \in \mathbf{G}(\mathbb{Q})_+$ and $k \in K$. In other words,

$$\varphi_\mathbf{H}(\sigma(Z_{x_2})) = [X_1^+ \times \{x_\sigma\}, h_{\sigma,\mathbb{Q}}]_K = [h_{\sigma,\mathbb{Q}}^{-1}(X_1^+ \times \{x_\sigma\}), 1]_K.$$

Since $\sigma(Z_{g,x_2})$ is an irreducible component of $T_{g^{-1}}(\varphi_\mathbf{H}(\sigma(Z_{x_2})))$, by Lemma 4.1,

$$\sigma(Z_{g,x_2}) = [g' h_{\sigma,\mathbb{Q}}^{-1}(X_1^+ \times \{x_\sigma\}), 1]_K$$

for some $g' \in \mathbf{G}(\mathbb{Q})_+$. Thus, letting $g_\sigma = g' h_{\sigma,\mathbb{Q}}^{-1}$, we can write $Z = Z_{g_\sigma, x_\sigma}$. $\square$

Remark 4.4. Using Remark 4.2, it is possible to augment the proof of Proposition 4.3 to show that, given a faithful representation $\mathbf{G} \to \mathbf{GL}_{m,\mathbb{Q}}$, the complexities of Z and $\sigma(Z)$ differ by at most a constant multiple independent of Z. However, the argument is more technical and also requires [Orr19, Theorem 1.3].

5. Height bound for representatives modulo congruence subgroups

We will need the following proposition on heights of certain elements in arithmetic groups. The $\mathbf{G} = \mathbf{SL}_2$ case of this proposition can be proved by the procedure outlined in [DS05, Exercise 1.2.2]. An effective version of the proposition for $\mathbf{G} = \mathbf{SO}_Q$ (where Q is an integral quadratic form) can be found at [LM16, Theorem 8].

Proposition 5.1. *Let $\mathbf{G}$ be a semisimple $\mathbb{Q}$-algebraic group and let $\Gamma \subset \mathbf{G}(\mathbb{Q})$ be a congruence subgroup. Let $\rho \colon \mathbf{G} \to \mathbf{GL}_{m,\mathbb{Q}}$ be a faithful representation. For each positive integer n, let*

$$\Gamma(n) = \Gamma \cap \rho^{-1}(\ker(\mathbf{GL}_m(\mathbb{Z}) \to \mathbf{GL}_m(\mathbb{Z}/n\mathbb{Z}))).$$

Then there exist constants C_{18} and C_{19} such that, for each positive integer n, every class in $\Gamma/\Gamma(n)$ has a representative in Γ whose 1-height is at most $C_{18} n^{C_{19}}$.

Our proof of Proposition 5.1 relies on the notion of expander families of graphs. An **expander family** is an infinite family of finite graphs (in which self-loops and multiple edges are permitted) for which there exists $\epsilon > 0$ such that for every graph G in the family,

$$\min\{\#\partial X/\#X : X \subset V(G), 0 < \#X \leq \#V(G)/2\} \geq \epsilon,$$

and there is a uniform upper bound for the degrees of the vertices of all graphs in the family. Here $V(G)$ means the set of vertices of G and ∂X denotes the set of edges of G which have one endpoint in X and the other endpoint not in X.

The relevance of expander families comes from the following theorem.

Theorem 5.2. *Let* $\mathbf{G}$ *be a semisimple* $\mathbb{Q}$-*algebraic group. Let* $\rho \colon \mathbf{G} \to \mathbf{GL}_{m,\mathbb{Q}}$ *be a faithful representation and let* $\Gamma = \rho^{-1}(\mathbf{GL}_m(\mathbb{Z}))$. *For each positive integer* n, *let* $\pi_n \colon \Gamma \to \mathbf{GL}_m(\mathbb{Z}/n\mathbb{Z})$ *denote the composition of* ρ *with reduction modulo* n.

Suppose that Γ *is infinite (equivalently,* $\mathbf{G}$ *is of non-compact type). Let* Δ *be a finite generating set for* Γ *such that* $\Delta^{-1} = \Delta$.

The Cayley graphs $\mathrm{Cay}(\pi_n(\Gamma), \pi_n(\Delta))$ *form an expander family as n ranges over the positive integers.*

Proof. Let $\tilde{\mathbf{G}}$ denote the simply connected cover of $\mathbf{G}$. If $\tilde{\mathbf{G}}$ is simple of real rank at least 2, then it has Kazhdan's property (T) [Vas68]. In general, Clozel [Clo03, Theorem 3.1] (using a result of Burger and Sarnak [BS91]) proved that the trivial representation is isolated in the support of the regular representation of $\tilde{\mathbf{G}}(\mathbb{R})$ on $L^2(\tilde{\mathbf{G}}(\mathbb{Q})\backslash\tilde{\mathbf{G}}(\mathbb{A}))$. Thanks to [LZ03, Proposition 4.7], this implies that Γ has property (τ) with respect to its congruence subgroups. By [LZ89, Proposition 1.2], this implies the theorem. $\square$

We will now use Theorem 5.2 to prove Proposition 5.1.

Proof of Proposition 5.1. The proposition is obvious if Γ is finite. Hence we may assume that Γ is infinite.

We begin by reducing to the case where $\Gamma = \rho^{-1}(\mathbf{GL}_m(\mathbb{Z}))$. Since Γ is a congruence subgroup, there exists n_0 such that

$$\rho^{-1}(\ker(\mathbf{GL}_m(\mathbb{Z}) \to \mathbf{GL}_m(\mathbb{Z}/n_0\mathbb{Z}))) = \Gamma(n_0) \subset \Gamma.$$

Since $\Gamma(n_0 n) \subset \Gamma(n)$, it suffices to prove the lemma with $\Gamma(n_0 n)$ in place of $\Gamma(n)$ (for any $n \in \mathbb{Z}$). Assuming that the proposition holds for $\rho^{-1}(\mathbf{GL}_m(\mathbb{Z}))$, we get a set of representatives for $\rho^{-1}(\mathbf{GL}_m(\mathbb{Z}))/\Gamma(n_0 n)$ of height bounded in terms of n. A subset of these representatives then form a set of representatives for $\Gamma(n_0)/\Gamma(n_0 n)$. Since $\Gamma(n_0)$ has finite index in Γ, we can choose once and for all a set of representatives $\gamma_1, \ldots, \gamma_s$ for $\Gamma/\Gamma(n_0)$. Then we can obtain representatives for every class in $\Gamma/\Gamma(n)$ by multiplying the γ_i by representatives for $\Gamma(n_0)/\Gamma(n_0 n)$.

Thus it suffices to assume that $\Gamma = \rho^{-1}(\mathbf{GL}_m(\mathbb{Z}))$.

According to [BHC62, Theorem 6.5], Γ is finitely generated, say by a set Δ. Enlarging Δ, we may assume that $\Delta^{-1} = \Delta$. Hence we can apply Theorem 5.2 to show that the Cayley graphs $\mathrm{Cay}_n = \mathrm{Cay}(\pi_n(\Gamma), \pi_n(\Delta))$ form an expander family.

By [Kow, Corollary 3.1.10], the graphs in the expander family $\{\mathrm{Cay}_n\}$ satisfy

$$\mathrm{diam}(\mathrm{Cay}_n) \ll \log(3 \,\#V(\mathrm{Cay}_n)).$$

Since $\#V(\mathrm{Cay}_n) = \#\pi_n(\Gamma) \leq n^{m^2}$ we deduce that

$$\mathrm{diam}(\mathrm{Cay}_n) \leq C_{20} \log(3n)$$

for some constant $C_{20} > 0$.

By the definition of the Cayley graph, any element $g \in \pi_n(\Gamma)$ can be written as the product of at most $\mathrm{diam}(\mathrm{Cay}_n)$ elements of $\pi_n(\Delta)$. Consequently we can find some $\tilde{g} \in \Gamma$ such that $\pi_n(\tilde{g}) = g$ and $\tilde{g}$ is a product of at most $\mathrm{diam}(\mathrm{Cay}_n)$ elements of Δ. Because $\mathrm{H}_1(xy) \leq m\mathrm{H}_1(x)\mathrm{H}_1(y)$ for all matrices x and y, we deduce that

$$\mathrm{H}_1(\tilde{g}) \leq m^{\mathrm{diam}(\mathrm{Cay}_n)-1} \max\{\mathrm{H}_1(\delta) : \delta \in \Delta\}^{\mathrm{diam}(\mathrm{Cay}_n)}$$

and thus

$$\mathrm{H}_1(\tilde{g}) \leq C_{21}^{C_{20} \log(3n)} = C_{22}\, n^{C_{23}}$$

for suitable constants. $\hfill\square$

Remark 5.3. The semisimplicity condition is necessary in Proposition 5.1: the proposition fails for the group $\mathbf{G} = \mathrm{Res}_{F/\mathbb{Q}}\, \mathbb{G}_m$, where F is a real quadratic field.

6. Parameter height bounds for Hecke–facteur families

The aim of this section is to prove versions of [DR18, Conjectures 12.2 and 12.7] for Hecke–facteur families, modified to use $\det^*(g)$ in place of $\deg(\pi(Y_{g,x_2}))$ (the same modification as we made to the definition of complexity).

Lemma 6.1. *Let* $(\mathbf{G}, X^+)$ *be a Shimura datum component and let* $S = \Gamma\backslash X^+$ *be an associated Shimura variety component. Fix a faithful representation* $\rho : \mathbf{G} \to \mathbf{GL}_{m,\mathbb{Q}}$. *Let* $(\mathbf{H}, X_{\mathbf{H}}^+, \mathbf{H}_1, \mathbf{H}_2)$ *be a facteur datum for* $(\mathbf{G}, X^+)$ *and choose fundamental sets* $\mathcal{F}_1 \subset X_1^+$, $\mathcal{F}_2 \subset X_2^+$ *and* $\mathcal{F} \subset X^+$ *as in Lemma 2.3.*

Then there exist constants C_{24} *and* C_{25} *such that, if* Y *is a pre-special subvariety in the Hecke–facteur family associated with* $(\mathbf{H}, X_{\mathbf{H}}^+, \mathbf{H}_1, \mathbf{H}_2)$ *satisfying* $Y \cap \mathcal{F} \neq \emptyset$, *then there exist* $g \in \mathbf{G}(\mathbb{Q})_+$ *and* $x_2 \in \mathcal{F}_2$ *satisfying* $Y = Y_{g,x_2}$, $\det^*(g) = N(\pi(Y))$ *and*

$$\mathrm{H}_1(g) \leq C_{24}\, N(\pi(Y))^{C_{25}}.$$

Proof. By the definition of the Hecke–facteur family, we can write

$$Y = Y_{g',x_2'} = g'(X_1^+ \times \{x_2'\})$$

for some $g' \in \mathbf{G}(\mathbb{Q})_+$ and $x_2' \in X_2^+$. We may choose g' so that $\det^*(g') = N(\pi(Y))$.

Pick a point $x \in Y \cap \mathcal{F}$ (which we are assuming to be non-empty). Then $g'^{-1}x \in X^+$ so by Lemma 2.3, we can choose

$$\gamma \in (\Gamma_1 \cap \mathbf{H}_1(\mathbb{R})_+).(\Gamma_2 \cap \mathbf{H}_2(\mathbb{R})_+) \subset \Gamma \cap \mathbf{H}(\mathbb{R})_+$$

such that $\gamma^{-1}g'^{-1}x \in \mathcal{F}_1 \times \mathcal{F}_2 \subset \mathcal{F}$.

Write $g = g'\gamma$ and let $g^{-1}x = (x_1, x_2)$. Here $x_i \in \mathcal{F}_i$. Because $\mathbf{H}_1$ commutes with $\mathbf{H}_2$, we have $Y = Y_{g,x_2}$.

Because x and $g^{-1}x$ are both in $\mathcal{F}$, [Orr18, Theorem 1.1] gives a polynomial bound for $\mathrm{H}_1(g)$ in terms of

$$|\det(g)| \cdot (\max\{b_{ij} : 1 \leq i,j \leq n\})^n$$

where the entries of g are written in lowest terms as a_{ij}/b_{ij} (with $b_{ij} > 0$). Since $\max\{b_{ij}\} \leq \mathrm{lcm}\{b_{ij}\}$, this implies that $H_1(g)$ is polynomially bounded in terms of $\det^*(g) = \det^*(g') = N(\pi(Y))$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

The following lemma is a modified special case of [DR18, Conjecture 12.7].

Proposition 6.2. *Let $(\mathbf{G}, X^+)$, S, ρ, $(\mathbf{H}, X_{\mathbf{H}}^+, \mathbf{H}_1, \mathbf{H}_2)$, $\mathcal{F}$, $\mathcal{F}_1$ and $\mathcal{F}_2$ be as in Lemma 6.1.*

Then there exist constants C_{26} and C_{27} such that, if Y is a pre-special subvariety in the Hecke–facteur family associated with $(\mathbf{H}, X_{\mathbf{H}}^+, \mathbf{H}_1, \mathbf{H}_2)$ satisfying $Y \cap \mathcal{F} \neq \emptyset$ and $z \in \mathcal{F} \cap \Gamma Y$, then there exists $\gamma \in \Gamma$ satisfying $z \in \gamma.Y$ and

$$H_1(\gamma) \leq C_{26} \, N(\pi(Y))^{C_{27}}.$$

Proof. Write $Y = Y_{g,x_2}$ as in Lemma 6.1.

We are given $z \in \mathcal{F} \cap \Gamma Y$, so we can pick $\gamma' \in \Gamma$ such that $z \in \gamma'Y$. Then

$$g^{-1}\gamma'^{-1}z \in X_1^+ \times \{x_2\}.$$

Because $\mathcal{F}_1$ is a fundamental set in X_1^+, we can choose $\gamma_1 \in \Gamma_1$ such that

$$\gamma_1^{-1}g^{-1}\gamma'^{-1}z \in \mathcal{F}_1 \times \{x_2\}.$$

Because we chose fundamental sets as in Lemma 2.3, $\mathcal{F}_1 \times \{x_2\} \subset \mathcal{F}$. Since also $z \in \mathcal{F}$, by [Orr18, Theorem 1.1] we get that $H_1(\gamma'g\gamma_1)$ is polynomially bounded in terms of $\det^*(\gamma'g\gamma_1) = \det^*(g) = N(\pi(Y))$. Because $\gamma', \gamma_1 \in \Gamma$, $\det^*(\gamma'g\gamma_1) = \det^*(g) = N(\pi(Y))$. Thus there are constants C_{28} and C_{29} such that

$$H_1(\gamma'g\gamma_1) \leq C_{28}N(\pi(Y))^{C_{29}}. \tag{3}$$

We can choose n, polynomially bounded in terms of $H_1(g)$, such that $\Gamma \cap g^{-1}\Gamma g$ contains the principal congruence subgroup $\Gamma(n)$ (defined with respect to ρ). Then $\Gamma(n) \cap \Gamma_1$ is a principal congruence subgroup in $\mathbf{H}_1(\mathbb{Q})$ (with respect to $\rho_{|\mathbf{H}_1}$). Applying Proposition 5.1 to $\mathbf{H}_1$, we can choose $\gamma_1' \in \Gamma_1$ such that $H_1(\gamma_1') \leq C_{18}n^{C_{19}}$ and $\gamma_1\gamma_1'^{-1} \in \Gamma(n) \cap \Gamma_1$.

To conclude, let

$$\gamma = \gamma'g\gamma_1\gamma_1'^{-1}g^{-1}.$$

Because $\gamma_1\gamma_1'^{-1} \in \Gamma(n) \subset g^{-1}\Gamma g$, we have $g\gamma_1\gamma_1'^{-1}g^{-1} \in \Gamma$ and hence $\gamma \in \Gamma$.

Since Γ_1 stabilises X_1^+ and acts trivially on X_2^+, we have

$$z \in \gamma'.Y = \gamma'g(X_1^+ \times \{x_2\}) = \gamma'g\gamma_1\gamma_1'^{-1}(X_1^+ \times \{x_2\}) = \gamma Y.$$

Finally, the height of $\gamma'g\gamma_1$ is polynomially bounded in terms of $N(\pi(Y))$ by (3). By definition, $H_1(\gamma_1')$ (and hence also $H_1(\gamma_1'^{-1})$) is polynomially bounded in terms of n and hence in terms of $H_1(g)$. By Lemma 6.1, $H_1(g)$ (and hence also $H_1(g^{-1})$) is polynomially bounded in terms of $N(\pi(Y))$. Combining these facts yields the bound for the height of γ. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

At last we are ready to prove the following special case of [DR18, Conjecture 12.2] (modified for our definition of complexity).

Proposition 6.3. *Let* $(\mathbf{G}, X^+)$, *S, ρ, $(\mathbf{H}, X_{\mathbf{H}}^+, \mathbf{H}_1, \mathbf{H}_2)$, $\mathcal{F}$, $\mathcal{F}_1$ and $\mathcal{F}_2$ be as in Lemma 6.1.*

Then there exist constants d, C_{30} and C_{31} such that, if Y is a pre-special sub-variety in the Hecke–facteur family associated with $(\mathbf{H}, X_{\mathbf{H}}^+, \mathbf{H}_1, \mathbf{H}_2)$ satisfying $Y \cap \mathcal{F} \neq \emptyset$, then there exists $g \in \mathbf{G}(\mathbb{Q})_+$ and a (pre-special) point $x \in X^+$ such that $Y = g\mathbf{H}_1(\mathbb{R})^+ g^{-1} x$ and

$$\mathrm{H}_d(g, x) \leq C_{30} \, \Delta(\pi(Y))^{C_{31}}.$$

Proof. As explained in [DO16, section 1.2], there is a constant d depending only on $(\mathbf{G}, X^+)$ such that the pre-special point x is defined over a number field of degree at most d. Because $g \in \mathbf{G}(\mathbb{Q})_+$, the height $\mathrm{H}_d(g, x)$ is finite for this value of d. If $(\mathbf{G}, X^+) = (\mathbf{GSp}_{2g}, \mathcal{H}_g)$, and ρ is the inclusion $\mathbf{GSp}_{2g} \hookrightarrow \mathbf{GL}_{2g}$, then the theory of complex multiplication of abelian varieties tells us that we can take $d = 2g$ (in particular, in the $E \times \mathrm{CM}$ case, we can take $d = 4$).

Write $Y = Y_{g,x_2}$ as in Lemma 6.1.

Let $x' \in Y$ denote a pre-special point such that $\pi(x')$ has minimal complexity in $\pi(Y)$. Then we may choose $\gamma' \in \Gamma$ such that $\gamma' x' \in \mathcal{F}$. By [DO16, Theorems 1.1 and 4.1], the height $\mathrm{H}_d(\gamma' x')$ is polynomially bounded in terms of $\Delta(\pi(x')) \leq \Delta(\pi(Y))$.

Since $\gamma' x' \in \mathcal{F} \cap \Gamma Y$, by Proposition 6.2, there exists $\gamma \in \Gamma$ with height polynomially bounded in terms of $N(\pi(Y)) \leq \Delta(\pi(Y))$, such that

$$\gamma' x' \in \gamma Y.$$

Therefore, $x = \gamma^{-1} \gamma' x'$ has height polynomially bounded in terms of $\Delta(\pi(Y))$. Meanwhile $x \in Y$ and so

$$Y = g\mathbf{H}_1(\mathbb{R})^+ g^{-1} x.$$

This concludes the proof. $\qquad\square$

7. Proof of Theorems 1.2 and 1.6

Theorem 1.2 is a special case of Theorem 1.6. The proof of Theorem 1.6 follows the same lines as that of [DR18, Theorem 14.2]. However, Theorem 1.6 is not directly a corollary of [DR18, Theorem 14.2] for three reasons:

(1) we have used a different definition of complexity;
(2) [DR18, Theorem 14.2] requires [DR18, Conjectures 11.1 and 12.2] to hold for all optimal points in V, while we have proved versions of these conjectures only for points in the $E \times \mathrm{CM}$ Hecke–facteur family;
(3) the proof of [DR18, Theorem 14.2] uses the property that $\Delta(Z) = \Delta(\sigma(Z))$, which holds for the definition of complexity used in [DR18] (see [Orr19, Theorem 1.3]) but not for our definition.

We will, therefore, outline the modifications to the proof of [DR18, Theorem 14.2] required in order to obtain Theorem 1.6. For the convenience of the reader, we recall the statement of Theorem 1.6 here.

Theorem. *(Theorem 1.6) Let $(\mathbf{G}, X^+)$ be a Shimura datum component and let $S = \Gamma \backslash X^+$ be an associated Shimura variety component. Let V be an irreducible algebraic curve in S not contained in any proper special subvariety.*

Let $(\mathbf{H}, X_{\mathbf{H}}^+, \mathbf{H}_1, \mathbf{H}_2)$ be a facteur datum in $(\mathbf{G}, X^+)$ and let $\mathfrak{F}$ be the associated Hecke–facteur family of special subvarieties of S. Suppose that, for any subvariety $Z \in \mathfrak{F}$, we have

$$\dim(Z) \leq \dim(S) - 2.$$

Let Σ be the set of points $s \in S$ for which the smallest special subvariety containing s is a member of $\mathfrak{F}$.

Assume that S, $\mathfrak{F}$, V and Σ satisfy Conjecture 1.7.

Then $V \cap \Sigma$ is finite.

Proof. Fix a faithful representation $\mathbf{G} \to \mathbf{GL}_{m,\mathbb{Q}}$ (which we use to define heights and complexities). Because we only consider points whose special closure is a member of the Hecke–facteur family $\mathfrak{F}$, the set Ω used in the proof of [DR18, Theorem 14.2] can be replaced by the one-element set $\{\mathbf{H}_1\}$. Let $\mathcal{F}$ be a fixed fundamental set as in Lemma 2.3. The constants d, $c_{\mathcal{F}}$ and $\delta_{\mathcal{F}}$ should be replaced by those afforded to us by Proposition 6.3 and the constants c_G and δ_G should be replaced by those afforded to us by Conjecture 1.7. As in [DR18], L is a finitely generated field of definition for V. We can and do assume that L contains the field F afforded to us by Proposition 4.3.

Consider a point $P \in V \cap \Sigma$, and let Z denote the smallest special subvariety of S containing P. By hypothesis, Z is a member of $\mathfrak{F}$. For each $\sigma \in \mathrm{Aut}(\mathbb{C}/L)$, let

$$z_\sigma \in \mathcal{V} \cap \pi^{-1}(\sigma(P)) \subset X^+,$$

where $\mathcal{V} := \pi^{-1}(V) \cap \mathcal{F}$. Let Y_σ denote the smallest pre-special subvariety of X^+ containing z_σ. Then, by Remark 7.1, $\pi(Y_\sigma) = \sigma(Z)$, and so, by Proposition 4.3, $\pi(Y_\sigma)$ belongs to $\mathfrak{F}$, for all $\sigma \in \mathrm{Aut}(\mathbb{C}/L)$.

By Proposition 6.3 (which replaces [DR18, Conjecture 12.2]),

$$Y_\sigma = g_\sigma \mathbf{H}_1(\mathbb{R})^+ g_\sigma^{-1} x_\sigma,$$

for $g_\sigma \in \mathbf{G}(\mathbb{Q})_+$ and $x_\sigma \in X^+$ with d-height polynomially bounded in terms of $\Delta(\pi(Y_\sigma)) = \Delta(\sigma(Z))$. By Conjecture 1.7 (which replaces [DR18, Conjecture 11.1]), $\Delta(\sigma(Z))$ is bounded above by a polynomial in

$$A := \# \mathrm{Aut}(\mathbb{C}/L) \cdot \sigma(P) = \# \mathrm{Aut}(\mathbb{C}/L) \cdot P.$$

Thus, the set Σ used in the proof of [DR18, Theorem 14.2] (which is not the set Σ used above) consists of tuples $(g_\sigma, x_\sigma, z_\sigma)$ which satisfy $H_d(g_\sigma, x_\sigma) \leq C_{32} A^{C_{33}}$. Since $\#\pi_2(\Sigma) = A$ (where π_2 projects on to the last co-ordinate), this is sufficient to apply the counting theorem [DR18, Theorem 9.1] (a variant of [HP16, Corollary 7.2]). The rest of the proof proceeds identically to that of [DR18, Theorem 14.2]. We just make two remarks regarding the arguments appearing on p.1879 of [DR18]:

(1) the equality $\dim(Y_1) = \dim\langle P_0 \rangle$ at line 20 is true, by Remark 7.1. In fact we need only the easier inequality $\dim(Y_1) \leq \dim(X) - 2$, which holds because Y_1 is a symmetric space for $F(\mathbb{R})^+$ and so has the same dimension as any member of $\mathfrak{F}$;

(2) the equality $\pi_2(F) = G_2$ at line –5 is not difficult to prove, but is non-trivial.

We conclude that A is bounded, so $\Delta(Z)$ is bounded. This implies the theorem. $\quad\square$

Remark 7.1. The arguments of [DR18] rely on the assertion appearing in that article (before Conjecture 11.1) that, for any $\sigma \in \mathrm{Aut}(\mathbb{C}/E_S)$ and any special subvariety Z of S, the subvariety $\sigma(Z)$ of S is special and $\Delta(\sigma(Z)) = \Delta(Z)$ (for the definition of Δ used in [DR18]). These assertions are true, though not proved in [DR18]. They are possibly known to the experts but, to our knowledge, the first proofs can be found in [Orr19].

Remark 7.2. Note that we removed the need for the equality $\Delta(\sigma(Z)) = \Delta(Z)$, used in [DR18, Theorem 14.2], or even any comparison between complexities of $\sigma(Z)$ and Z, by comparing $H_d(g_\sigma, x_\sigma)$ directly with $\# \mathrm{Aut}(\mathbb{C}/L) \cdot \sigma(P)$, which is trivially equal to $\# \mathrm{Aut}(\mathbb{C}/L) \cdot P$.

8. André's height bound

The key step in our proof of a special case of the large Galois orbits conjecture, as required for Theorem 1.1, is a generalisation of a height bound due to André for fibres with large endomorphism rings in an abelian scheme over a curve with a point of purely multiplicative reduction. This height bound is stated and proved in [And89, Ch. X] for abelian schemes of odd relative dimension g. We want to apply it for $g = 2$, so we need a generalisation of the theorem. It turns out that we can replace the condition that g is odd by the condition that the generic endomorphism algebra of the abelian scheme is a totally real field of odd degree.

The theorem is a height bound for fibres of the abelian scheme whose endomorphism algebra is large in the following sense.

Definition. An abelian variety A of dimension g is **exceptional** if there is no injection from $\mathrm{End}(A)$ to $\mathrm{M}_g(\mathbb{Q})$.

Theorem 8.1. *Let V' be a smooth connected (not necessarily proper) curve over a number field K. Let s_0 be a closed point in V' and let $V = V' \setminus \{s_0\}$. Let $\bar{\eta}$ be a geometric generic point of V. Let h denote a Weil height on V'.*

Let $\mathfrak{A} \to V$ be an abelian scheme with purely multiplicative reduction at s_0 (that is, if $\mathfrak{A}' \to V'$ is the connected Néron model of $\mathfrak{A}$ over V' then the fibre $\mathfrak{A}'_{s_0}$ is a torus). Let g be the relative dimension of $\mathfrak{A} \to V$.

Assume that:

(i) $g > 1$.

(ii) $\mathrm{End}(\mathfrak{A}_{\bar{\eta}}) \otimes \mathbb{Q}$ is a totally real field E.

(iii) $[E : \mathbb{Q}]$ is odd.

(iv) The generic Mumford–Tate group of $\mathfrak{A}$ is $\mathbb{G}_{m,\mathbb{Q}} . \operatorname{Res}_{E/\mathbb{Q}} \operatorname{Sp}_{2n,E}$, where $n = g/[E : \mathbb{Q}]$.

Then there exist constants C_{34} and C_{35} such that, for every point $s \in V(\overline{\mathbb{Q}})$, if $\mathfrak{A}_s$ is exceptional then

$$h(s) \leq C_{34} \left[K(s) : K \right]^{C_{35}}.$$

Remark 8.2. Theorem 8.1 can be found at [And89, Ch. X, Theorem 1.3] with conditions (ii)–(iv) replaced by: g is odd and $\mathfrak{A}_{\bar{\eta}}$ is simple. These conditions, together with the multiplicative reduction condition, imply (ii)–(iv): (ii) is [And89, Ch. X, Lemma 2.2]; (iii) holds because the degree of a totally real field acting on a simple abelian variety A must divide $\dim(A)$ [Mum74, Ch. IV, §20, p. 191, Corollary]; (iv) is [And89, Ch. X, 3.3 Sublemma]. On the other hand, when g is even, simplicity of $\mathfrak{A}_{\bar{\eta}}$ is not sufficient to imply conditions (ii)–(iv). When $g = 2$, condition (iv) is automatic because of the classification in section 2.F. However, when $g = 4$, condition (iv) is famously not automatic, as first observed by Mumford.

Remark 8.3. Theorem 8.1 is certainly not as general as possible. In particular, condition (iii) can be relaxed. For example, Masser has recently obtained a similar theorem for abelian schemes whose generic fibre is a principally polarised, simple abelian surface with real multiplication [Mas].

Remark 8.4. Observe that, if the abelian scheme $\mathfrak{A} \to V$ is principally polarised and the image of the associated map $V \to \mathcal{A}_g$ is Hodge generic, then $\operatorname{End}(\mathfrak{A}_{\bar{\eta}}) = \mathbb{Z}$ and conditions (ii)–(iv) are satisfied.

Remark 8.5. For $g = 2$, abelian surfaces of $E \times$ CM type are exceptional with respect to $\mathbb{Q}$: their endomorphism algebra is $\mathbb{Q} \times F$, $F_1 \times F_2$ or $\mathrm{M}_2(F)$ where F, F_1, F_2 are imaginary quadratic fields. Each of these algebras contains a commutative subalgebra of dimension at least 3 over $\mathbb{Q}$, so they do not embed into $\mathrm{M}_2(\mathbb{Q})$. Abelian surfaces whose endomorphism algebra is a non-split quaternion algebra over $\mathbb{Q}$ are also exceptional. However abelian surfaces isogenous to the square of a non-CM elliptic curve are not exceptional (their endomorphism algebra is $\mathrm{M}_2(\mathbb{Q})$).

Remark 8.6. We can replace the "multiplicative reduction" condition in Theorem 8.1 by the condition: there exists a semiabelian scheme $\mathfrak{A}' \to V'$ such that $\mathfrak{A}'_{|V} \cong \mathfrak{A}$ and $\mathfrak{A}'_{s_0}$ is a torus. This is because [BLR90, Ch. 7, Prop. 3] tells us that, if $\mathfrak{A}'$ is a semiabelian scheme whose generic fibre is an abelian variety, then $\mathfrak{A}'$ is its own connected Néron model.

Before explaining how to modify the proof of [And89, Ch. X, Theorem 1.3] to obtain Theorem 8.1, we prove the properties of endomorphism algebras of exceptional abelian varieties which we will use.

Lemma 8.7. *Let D be a quaternion algebra over a totally real field E. (We allow D to be split.) Let $\dagger$ be a positive involution of the $\mathbb{Q}$-algebra D. Then there exists $\alpha \in D^\times$ such that $\alpha^\dagger = -\alpha$ and $E(\alpha)$ is a CM field.*

Proof. First suppose that D is non-split over E.

A positive involution must act trivially on the totally real field E, so † is E-linear. Since it is an involution, it is diagonalisable with eigenvalues ± 1. A positive involution always acts non-trivially on a quaternion algebra, so we deduce that

$$\{\alpha \in D : \alpha^\dagger = -\alpha\}$$

is a non-zero E-linear subspace of D.

Choose any non-zero element α of this subspace. The subalgebra $E(\alpha) \subset D$ is a field because D is a division algebra. Because $\alpha^\dagger = -\alpha$, † restricts to a non-trivial involution of the field $E(\alpha)$. This restriction remains a positive involution, and it is well-known that any field with a non-trivial positive involution is a CM field.

Now consider the case where D is split. We may choose an isomorphism $D \to \mathrm{M}_2(E)$ under which † becomes either (a) matrix transposition or (b) the involution $A \mapsto (\det A)A^{-1}$. In either case, $\alpha = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ satisfies $\alpha^\dagger = -\alpha$. Since $\alpha^2 = -1$, the algebra $E(\alpha) \subset D$ is an extension of the totally real field E obtained by adjoining a square root of a totally negative element of E. Thus $E(\alpha)$ is a CM field. $\qquad \square$

Lemma 8.8. *Let A be an abelian variety of dimension g. Let $\hat{E}$ be a maximal commutative subalgebra of $\mathrm{End}(A) \otimes \mathbb{Q}$. If A is exceptional, then $\hat{E}$ is not a field of odd degree over $\mathbb{Q}$.*

Proof. Assume for contradiction that $\hat{E}$ is a field of odd degree. Since $\mathrm{End}(A) \otimes \mathbb{Q}$ contains a maximal commutative subalgebra which is a field, $\mathrm{End}(A) \otimes \mathbb{Q}$ must be a simple algebra so it is isomorphic to $\mathrm{M}_r(D)$ for some division algebra D and positive integer r.

We split into cases depending on the type of D in the Albert classification.

If D has type I (a totally real field), then $[\hat{E} : \mathbb{Q}] = r[D : \mathbb{Q}]$. But $[\hat{E} : \mathbb{Q}]$ divides g. Hence, $\mathrm{M}_r(D)$ injects into $\mathrm{M}_g(\mathbb{Q})$, contradicting the exceptionality of A.

If D has type II or III (a quaternion algebra) or type IV (a division algebra whose centre is a CM field), then any maximal commutative subalgebra of $\mathrm{M}_r(D)$ has even degree over $\mathbb{Q}$, contradicting our hypothesis. $\qquad \square$

Lemma 8.9. *Let E be a totally real field of odd degree. Let A be a polarised abelian variety of dimension g, equipped with an embedding $E \to \mathrm{End}(A) \otimes \mathbb{Q}$ whose image is stabilised by the Rosati involution. If A is exceptional and the degree of a maximal commutative subalgebra of $\mathrm{End}(A) \otimes \mathbb{Q}$ is $2[E : \mathbb{Q}]$, then there exists a maximal commutative subalgebra of $\mathrm{End}(A) \otimes \mathbb{Q}$ which contains E, is a CM field, and is stabilised by the Rosati involution.*

Proof. First we show that A is isotypic. Suppose not. Then $\mathrm{End}(A) \otimes \mathbb{Q}$ is not simple, so its maximal commutative subalgebras are never fields. Since A has a maximal commutative subalgebra of degree 2 over E, it must have $E \times E$ as a maximal commutative subalgebra. Therefore, A is isogenous to a product $A_1 \times A_2$, where $\mathrm{End}(A_1) \otimes \mathbb{Q}$ and $\mathrm{End}(A_2) \otimes \mathbb{Q}$ both have E as a maximal commutative

subalgebra. By Lemma 8.8, A_1 and A_2 are not exceptional. In other words, $\mathrm{End}(A_1) \otimes \mathbb{Q}$ injects into $\mathrm{M}_{g_1}(\mathbb{Q})$ and $\mathrm{End}(A_2) \otimes \mathbb{Q}$ injects into $\mathrm{M}_{g_2}(\mathbb{Q})$, where $g_i = \dim(A_i)$. Since $\mathrm{End}(A) \otimes \mathbb{Q} \cong (\mathrm{End}(A_1) \otimes \mathbb{Q}) \times (\mathrm{End}(A_2) \otimes \mathbb{Q})$, we deduce that $\mathrm{End}(A) \otimes \mathbb{Q}$ injects into $\mathrm{M}_{g_1+g_2}(\mathbb{Q})$. This contradicts the exceptionality of A.

Thus A is isotypic. Let Z denote the centraliser of E in $\mathrm{End}(A) \otimes \mathbb{Q}$. Since E is commutative, $E \subset Z$ and so by the double centraliser theorem, E is the centre of Z. Any maximal commutative subalgebra of $\mathrm{End}(A) \otimes \mathbb{Q}$ containing E is also a maximal commutative subalgebra of Z; since such a maximal commutative subalgebra has degree 2 over E, we deduce that Z is a quaternion algebra over E (perhaps a split quaternion algebra).

Let $\dagger$ denote the Rosati involution on $\mathrm{End}(A) \otimes \mathbb{Q}$. Since $\dagger$ stabilises E, it also stabilises its centraliser Z. Applying Lemma 8.7, we get $\alpha \in Z^{\times}$ such that $\alpha^{\dagger} = -\alpha$ and $E(\alpha)$ is a CM field. Now $[E(\alpha) : \mathbb{Q}] = 2[E : \mathbb{Q}]$ is a maximal commutative subalgebra of $\mathrm{End}(A) \otimes \mathbb{Q}$. Furthermore the fact that $\alpha^{\dagger} = -\alpha$ implies that $\dagger$ stabilises $E(\alpha)$. Thus $E(\alpha)$ is the desired maximal commutative subalgebra. $\qquad\square$

We now describe the modifications to the proof of [And89, Ch. X, Theorem 1.3] required to obtain the more general Theorem 8.1.

Proof of Theorem 8.1. As mentioned above, conditions (ii) and (iv) in Theorem 8.1 are precisely the conclusions of [And89, Ch. X, Lemma 2.2 and Sublemma 3.3]. Other than in these lemmas, the only place where [And89] uses the oddness of g is in Construction 2.4.1, so this is the only part of the proof which we need to modify.

First we recall some of the notation from [And89, Ch. X]. Let s be a point in $V(\overline{\mathbb{Q}})$ such that $\mathfrak{A}_s$ is exceptional. Let $E = \mathrm{End}(\mathfrak{A}_{\bar{\eta}}) \otimes \mathbb{Q}$ (we are assuming that this is a totally real field of odd degree) and let $\hat{E}$ denote a maximal commutative subalgebra of $\mathrm{End}(\mathfrak{A}_s) \otimes \mathbb{Q}$ which contains E. Let $n = g/[E : \mathbb{Q}]$.

Let $f \colon \mathfrak{A} \to V$ denote the structure map of our abelian scheme. Let F be the Galois closure of E and let $\hat{F}$ be the compositum of the Galois closures of the simple factors of $\hat{E}$. Then the local system of F-vector spaces $R_1 f_{\mathbb{C},*}^{\mathrm{an}}(F)$ splits as a direct sum $\bigoplus_{\sigma \colon E \to \mathbb{C}} W_{\sigma}$, where E acts on W_{σ} via σ. For each σ, W_{σ} has F-rank $2n$. Similarly $H_1(\mathfrak{A}_s(\mathbb{C}), \hat{F})$ splits as a direct sum $\bigoplus_{\hat{\sigma} \colon \hat{E} \to \mathbb{C}} \hat{W}_{\hat{\sigma}}$. These splittings are compatible in the sense that

$$(W_{\sigma})_s \otimes_F \hat{F} = \bigoplus_{\hat{\sigma} \mid \sigma} \hat{W}_{\hat{\sigma}}.$$

Replacing K by a finite extension and V by a Zariski open subset of an étale cover, we may assume that $F \subset K$, $\mathrm{End}(\mathfrak{A}_{\eta}) = \mathrm{End}(\mathfrak{A}_{\bar{\eta}})$ (where η denotes the generic point of V as a K-variety) and the $\mathcal{O}_V$-module $H_{DR}^1(\mathfrak{A}/V)$ is free. Let $\hat{K}$ denote the compositum $K(s)\hat{F}$.

Similarly to the above, we get splittings

$$H^0(V, H^1_{DR}(\mathfrak{A}/V)) \otimes_{\mathcal{O}(V)} K(V) = \bigoplus_{\sigma \colon E \to K(V)} W^\sigma_{DR},$$

$$H^1_{DR}(\mathfrak{A}_s/\hat{K}) = \bigoplus_{\hat{\sigma} \colon \hat{E} \to \hat{K}} \hat{W}^{\hat{\sigma}}_{DR}.$$

Choose a polarisation of the abelian scheme $\mathfrak{A} \to V$. Thanks to the isomorphism

$$R^2 f^{\mathrm{an}}_{\mathbb{C},*}(\mathbb{Q}(1)) \cong \mathrm{Hom}\left(\bigwedge\nolimits^2 R_1 f^{\mathrm{an}}_{\mathbb{C},*}(\mathbb{Z}), \mathbb{Q}(1)\right),$$

this polarisation gives rise to a symplectic form $2\pi i \langle, \rangle_\mathbb{Q}$ on $R_1 f^{\mathrm{an}}_{\mathbb{C},*}(\mathbb{Q})$ with values in $\mathbb{Q}(1)$. Because E is totally real, this form (after extension of scalars to F) is a sum of symplectic forms on the components W_σ which we denote by

$$2\pi i \langle, \rangle_\sigma \colon W_\sigma \times W_\sigma \to F(1).$$

Let Δ be the image of a holomorphic embedding of the unit disc in $V'(\mathbb{C})$ mapping 0 to s_0, let $\Delta^* = \Delta \setminus \{s_0\}$ and let U be a simply connected dense open subset of Δ^*. Note that we can cover $V(\mathbb{C})$ by finitely many such sets U, so we may assume that our exceptional point s is in U.

Let W^1_σ denote the maximal constant subsystem of $W_{\sigma|\Delta^*}$. As a consequence of the multiplicative reduction condition at s_0, W^1_σ is totally isotropic with respect to $\langle, \rangle_\sigma$ and has F-rank n [And89, Ch. X, Lemma 2.3].

Choose bases $\gamma_{\tau,1}, \ldots, \gamma_{\tau,n}$ for W^1_τ and $\omega_{\sigma,1}, \ldots, \omega_{\sigma,2n}$ for W^σ_{DR}. The "locally invariant periods" of $\mathfrak{A}/V$ are the meromorphic functions on Δ^* defined by

$$t \mapsto \int_{\gamma_{\tau,j}(t)} \omega_{\sigma,i}(t)$$

(for $\sigma \colon E \to K(V)$, $\tau \colon E \to \mathbb{C}$, $1 \le i \le 2n$, $1 \le j \le n$). These are the G-functions which lie at the heart of the proof of Theorem 8.1.

We now define symplectic forms $\langle, \rangle^\sigma_{DR} \colon W^\sigma_{DR} \times W^\sigma_{DR} \to \hat{K}$ which are dual to the forms $2\pi i \langle, \rangle_\sigma$ in a sense which we will make precise below.

Let ξ denote the class in $H^2_{DR}(\mathfrak{A}_\eta/K(V))$ induced by our chosen polarisation of $\mathfrak{A}$. Since $\dim_{K(V)} H^{2g}_{DR}(\mathfrak{A}_\eta/K(V)) = 1$ and $\xi^g \ne 0$, the following formula defines a symplectic form $\langle, \rangle_{DR}$ on $H^1_{DR}(\mathfrak{A}_\eta/K(V))$ with values in $K(V)$:

$$\alpha \cup \beta \cup \xi^{g-1} = \frac{1}{g} \langle \alpha, \beta \rangle_{DR} \, \xi^g \in H^{2g}_{DR}(\mathfrak{A}_\eta/K(V)).$$

Let $\mathcal{M}(U)$ denote the field of meromorphic functions on U. Write $R_1 f_*(\mathbb{Q}(1)) \otimes_\mathbb{Q} \mathcal{M}(U)$ as short hand for $H^0(U, R_1 f^{\mathrm{an}}_{\mathbb{C},*}(\mathbb{Q}(1))) \otimes_\mathbb{Q} \mathcal{M}(U)$. The local system $W_{\sigma|U}$ is trivial because U is simply connected, so the pairing $(\gamma, \omega) \mapsto \int_\gamma \omega$ (in each fibre of $\mathfrak{A}_{|U} \to U$) induces an isomorphism

$$R_1 f_*(\mathbb{Q}(1)) \otimes_\mathbb{Q} \mathcal{M}(U) \to (H^1_{DR}(\mathfrak{A}_\eta/K(V)) \otimes_{K(V)} \mathcal{M}(U))^\vee. \tag{4}$$

Under this isomorphism, the symplectic form $2\pi i \langle,\rangle_{\mathbb{Q}} \in \wedge^2 (R_1 f_*(\mathbb{Q}(1)) \otimes_{\mathbb{Q}} \mathcal{M}(U))^\vee$ corresponds to $\xi \in H^2_{DR}(\mathfrak{A}_\eta/K(V)) = \wedge^2 H^1_{DR}(\mathfrak{A}_\eta/K(V))$. A calculation shows that, if $e_1, \ldots, e_{2g}$ is a symplectic basis for $R_1 f^{an}_{\mathbb{C},*}(\mathbb{Q}(1))_{|U}$ with respect to $2\pi i \langle,\rangle_{\mathbb{Q}}$ and $e_1^\vee, \ldots, e_{2g}^\vee$ is the corresponding dual basis for $H^1_{DR}(\mathfrak{A}_\eta/K(V)) \otimes_{K(V)} \mathcal{M}(U)$, then $e_1^\vee, \ldots, e_{2g}^\vee$ is a symplectic basis with respect to $\langle,\rangle_{DR}$.

Since $2\pi i \langle,\rangle_{\mathbb{Q}}$ is the orthogonal sum of symplectic forms on W_σ and the isomorphism (4) is compatible with the actions of E, we deduce that $\langle,\rangle_{DR}$ is the orthogonal sum of symplectic forms

$$\langle,\rangle^\sigma_{DR} \colon W^\sigma_{DR} \times W^\sigma_{DR} \to \hat{K}.$$

The notational set-up is complete. We now explain the modifications of [And89, Ch. X, Construction 2.4.1] needed to obtain non-trivial global relations between the values of the locally invariant periods at a point $s \in V(\overline{\mathbb{Q}})$ where $\mathfrak{A}_s$ is exceptional.

The initial remarks of [And89, Ch. X, Construction 2.4.1] show that $\hat{E} \neq E$. The argument in [And89] relies on g being odd, but we can replace this by Lemma 8.8.

If $[\hat{E} : E] \geq 3$, then we are in Case 1 of [And89, Ch. X, Construction 2.4.1]: for each $\sigma \colon E \to \mathbb{C}$ there are at least three $\hat{\tau} \colon \hat{E} \to \mathbb{C}$ extending σ. Hence at least one of these $\hat{\tau}$ must satisfy $\dim_{\hat{F}}(\hat{W}_{\hat{\tau}}) < \frac{1}{2} \dim_F((W_\sigma)_s)$. The proof for Case 1 in [And89] does not use the oddness of g, so it works without change.

If $[\hat{E} : E] = 2$, then we may choose $\hat{E}$ according to Lemma 8.9. For $\sigma \colon E \to \mathbb{C}$, let $\hat{\sigma}_1$ and $\hat{\sigma}_2$ denote the two homomorphisms $\hat{E} \to \mathbb{C}$ extending σ. An element of $\text{Aut}(\mathbb{C})$ which exchanges $\hat{\sigma}_1$ and $\hat{\sigma}_2$ will exchange $\hat{W}_{\hat{\sigma}_1}$ and $\hat{W}_{\hat{\sigma}_2}$ so

$$\dim_{\hat{F}}(\hat{W}_{\hat{\sigma}_1}) = \dim_{\hat{F}}(\hat{W}_{\hat{\sigma}_2}) = \tfrac{1}{2} \dim_F((W_\sigma)_s).$$

Thus we are in either Case 2 or Case 3 of [And89, Ch. X, Construction 2.4.1].

Case 2 occurs when $\hat{W}_{\hat{\sigma}_2} \cap \left((W^1_\sigma)_s \otimes_F \hat{F}\right) \neq 0$. Again the proof in [And89] does not use the oddness of g, so no change is required.

Finally we consider Case 3, where $\hat{W}_{\hat{\sigma}_2} \cap \left((W^1_\sigma)_s \otimes_F \hat{F}\right) = 0$. In the setting of [And89], the oddness of g implies that $\hat{E}$ is a CM field. In our more general setting, it might not be true that every choice of maximal commutative subalgebra $\hat{E} \subset \text{End}(\mathfrak{A}_s) \otimes \mathbb{Q}$ containing E is a CM field, but Lemma 8.9 ensures that it is possible to choose an $\hat{E}$ which is a CM field.

The fact that $\hat{E}$ is a CM field stabilised by the Rosati involution is crucial for the construction of period relations. Indeed, the Rosati involution restricts to complex conjugation on $\hat{E}$ and so $\hat{W}_{\hat{\sigma}_2}$ is isotropic with respect to $\langle,\rangle_\sigma$. This ensures that we can choose a basis $\gamma_{\sigma,n+1}, \ldots, \gamma_{\sigma,2n}$ for $\hat{W}_{\hat{\sigma}_2}$ which, when combined with the previously chosen basis $\gamma_{\sigma,1}(s), \ldots, \gamma_{\sigma,n}(s)$ for $(W^1_\sigma)_s$, forms a symplectic basis for $(W_\sigma)_s \otimes_F \hat{F}$. Similarly, the fact that $\hat{E}$ is a CM field stabilised by the Rosati involution implies that we can choose a symplectic basis $\hat{\omega}_{\sigma,1}, \ldots, \hat{\omega}_{\sigma,2n}$ for W^σ_{DR} in such a way that $\hat{\omega}_{\sigma,1}(s), \ldots, \hat{\omega}_{\sigma,n}(s) \in \hat{W}^{\hat{\sigma}_1}_{DR}$ and $\hat{\omega}_{\sigma,n+1}(s), \ldots, \hat{\omega}_{\sigma,2n}(s) \in \hat{W}^{\hat{\sigma}_2}_{DR}$.

Because these bases for $(W_\sigma)_s \otimes_F \hat{F}$ and W_{DR}^σ are symplectic, the resulting period matrix satisfies the Riemann relations [And89, Ch. X, (2.3.1)–(2.3.3)]. The remainder of [And89, Ch. X, Construction 2.4.1, Case 3], using these relations to obtain a quadratic relation between locally invariant periods at s, does not use the oddness of g so we can reuse it without change. $\qquad\square$

9. Galois orbits

In this section we prove special cases of our Galois orbits conjecture and deduce Theorems 1.1 and 1.4. The proof of Theorem 1.4 combines Theorem 8.1 with the Masser–Wüstholz isogeny theorem and a refinement of the Brauer–Siegel theorem. To deduce Theorem 1.1 from Theorem 1.4, we use a toroidal compactification of $\mathcal{A}_g$ to construct a semiabelian scheme to which we can apply Theorem 1.4. We also show that the two formulations of the Galois orbits conjecture, Conjectures 1.3 and 1.5 are equivalent.

Given an abelian surface scheme $\mathfrak{A} \to V$, recall that we defined $N(s)$ and $\Delta'(s)$ for $E \times$ CM points $s \in V$ in Remark 3.3.

Proposition 9.1. *Conjectures 1.3 and 1.5 are equivalent.*

Proof. Firstly let $\mathfrak{A} \to V$ be an abelian scheme as in Conjecture 1.5, defined over a finitely generated field $L \subset \mathbb{C}$. This abelian scheme induces a morphism of varieties $q \colon V \to \mathcal{A}_2$, also defined over L. Since $\mathfrak{A} \to V$ is not isotrivial, the image $q(V)$ is still a curve. Since $\operatorname{End}(\mathfrak{A}_{\bar\eta}) = \mathbb{Z}$, $q(V)$ is not contained in any proper special subvariety of $\mathcal{A}_2$ (see section 2.F). A point $s \in V$ is an $E \times$ CM point if and only if $q(s) \in \Sigma$. Hence we can apply Conjecture 1.3 to the Zariski closure of $q(V)$ in $\mathcal{A}_2$, obtaining a bound
$$\# \operatorname{Aut}(\mathbb{C}/L) \cdot q(s) \geq C_{36} \Delta(Z)^{C_{37}}$$
for all $E \times$ CM points $s \in V$, where Z is the unique $E \times$ CM curve containing $q(s)$. Using Lemma 3.6 and the fact that $\# \operatorname{Aut}(\mathbb{C}/L) \cdot s \geq \# \operatorname{Aut}(\mathbb{C}/L) \cdot q(s)$, we deduce that Conjecture 1.3 holds.

Conversely, let $V \subset \mathcal{A}_2$ be an algebraic curve as in Conjecture 1.3, defined over a finitely generated field $L \subset \mathbb{C}$. Let $\tilde{V}$ be an irreducible component of the preimage of V in $\mathcal{A}_{2,3}$. This is defined over a finite extension $\tilde{L}$ of L. The universal abelian scheme over $\mathcal{A}_{2,3}$ restricts to an abelian scheme $\tilde{\mathfrak{A}} \to \tilde{V}$. Since V is not contained in any special subvariety of $\mathcal{A}_2$, $\operatorname{End}(\tilde{\mathfrak{A}}_{\bar\eta}) = \mathbb{Z}$. If $s \in V \cap \Sigma$, then we can find an $E \times$ CM point $\tilde{s} \in \tilde{V}$ such that $q(\tilde{s}) = s$. Thanks to Conjecture 1.5, this satisfies
$$\# \operatorname{Aut}(\mathbb{C}/\tilde{L}) \cdot \tilde{s} \geq C_{38} \Delta'(\tilde{s})^{C_{39}}.$$
Because q is a finite morphism,
$$\# \operatorname{Aut}(\mathbb{C}/L) \cdot s \geq C_{40} \# \operatorname{Aut}(\mathbb{C}/\tilde{L}) \cdot \tilde{s}.$$
We can therefore use Lemma 3.6 to complete the proof of Conjecture 1.5. $\qquad\square$

As a first step towards the proof of Conjecture 1.5 under the conditions of Theorem 1.4, we prove a Galois bound relative to $N(s)$. We then combine this with a lower bound for class numbers of imaginary quadratic fields.

Proposition 9.2. *Let V be an irreducible algebraic curve over $\overline{\mathbb{Q}}$ and let $\mathfrak{A} \to V$ be a principally polarised non-isotrivial abelian scheme of relative dimension 2. Suppose that $\mathrm{End}(\mathfrak{A}_{\overline{\eta}}) = \mathbb{Z}$, where $\overline{\eta}$ denotes a geometric generic point of V.*

Suppose also that there exist a curve V', a semiabelian scheme $\mathfrak{A}' \to V'$ and an open immersion $\iota\colon V \to V'$ (all defined over $\overline{\mathbb{Q}}$) such that $\mathfrak{A} \cong \iota^* \mathfrak{A}'$ and there is some point $s_0 \in V'(\overline{\mathbb{Q}})$ for which the fibre $\mathfrak{A}'_{s_0}$ is a torus.

Let L be a number field over which V and $\mathfrak{A} \to V$ are defined.

Then there exist positive constants C_{41} and C_{42} such that, for every $E \times$ CM point $s \in V$,

$$[L(s) : L] \geq C_{41}\, N(s)^{C_{42}}.$$

Proof. After a finite extension of L, we may assume that V', $\mathfrak{A}' \to V'$, $\iota\colon V \to V'$ and s_0 are all defined over L.

Because $\mathrm{End}(\mathfrak{A}_{\overline{\eta}}) = \mathbb{Z}$ and $\dim(\mathfrak{A}_{\overline{\eta}}) = 2$, the Mumford–Tate group of $\mathfrak{A}_{\overline{\eta}}$ is $\mathbf{GSp}_{4,\mathbb{Q}}$ (see section 2.F). Thus $\mathfrak{A} \to V$ satisfies the conditions of Theorem 8.1, as modified in Remark 8.6.

Let s be an $E \times$ CM point in V. The image of s under the map $V \to \mathcal{A}_2$ induced by $\mathfrak{A} \to V$ is in the intersection between the image of V and a special curve. Since the image of V is a curve defined over a number field, we deduce that $s \in V(\overline{\mathbb{Q}})$.

Now $\mathrm{End}(\mathfrak{A}_s) \otimes \mathbb{Q}$ contains $\mathbb{Q} \times F$ where F is an imaginary quadratic field. This is a commutative $\mathbb{Q}$-algebra of degree 3, so cannot inject into $\mathrm{M}_2(\mathbb{Q})$. Hence $\mathfrak{A}_s$ is an exceptional abelian surface in the sense of Theorem 8.1.

Therefore by Theorem 8.1, $h(s)$ is polynomially bounded in terms of $[L(s) : L]$, where h denotes a Weil height on V'. Let h_F denote the stable Faltings height. As proved in [Fal83, p. 356],

$$|h_F(\mathfrak{A}_s) - h(s)| = \mathrm{O}(\log h(s)).$$

We conclude that $h_F(\mathfrak{A}_s)$ is polynomially bounded in terms of $[L(s) : L]$.

By Lemma 3.1, there exist elliptic curves E_1 and E_2 and a polarised isogeny $E_1 \times E_2 \to \mathfrak{A}_s$ of degree $N(s)$, and $N(s)$ is the minimum degree of any polarised isogeny $E_1 \times E_2 \to \mathfrak{A}_s$. Thanks to Lemma 2.1, $N(s)$ is the minimum degree of any isogeny $E_1 \times E_2 \to \mathfrak{A}_s$ (polarised or not). By Lemma 3.2, E_1 and E_2 are the only pair of elliptic curves whose product possesses an isogeny to $\mathfrak{A}_s$ of degree $N(s)$. Therefore, for every $\sigma \in \mathrm{Aut}(\mathbb{C}/L(s))$, $E_1^\sigma \cong E_1$ and $E_2^\sigma \cong E_2$ (E_2 has CM and E_1 does not, so σ cannot swap them). Since elliptic curves are defined over their fields of moduli, we conclude that E_1 and E_2 are both defined over $L(s)$.

Therefore by the Masser–Wüstholz isogeny theorem [MW93], there exists an isogeny $E_1 \times E_2 \to \mathfrak{A}_s$ of degree at most

$$C_{43}\, \max(h_F(\mathfrak{A}_s), [L(s):L])^{C_{44}}.$$

We conclude that $N(s)$ also satisfies this bound. Combining this with the fact that $h_F(\mathfrak{A}_s)$ is polynomially bounded in terms of $[L(s) : L]$ completes the proof. $\qquad\square$

Proposition 9.3. *Let $\mathfrak{A} \to V$ be an abelian scheme satisfying the conditions of Proposition 9.2 (including the multiplicative reduction condition).*
Then Conjecture 1.5 holds for $\mathfrak{A} \to V$.

Proof. Let L be a number field over which V and $\mathfrak{A} \to V$ are defined. Let $s \in V$ be an $E \times$ CM point. By Lemma 3.1, there exist elliptic curves E_1 and E_2 and an isogeny $E_1 \times E_2 \to A_s$ of degree $N(s)$, where E_2 has CM and E_1 does not. Then

$$\Delta'(s) = \max\{N(s), |\mathrm{disc}(\mathrm{End}(E_2))|\}.$$

According to the argument from the proof of Proposition 9.2, E_1 and E_2 are defined over $L(s)$. Therefore by [Len87, Proposition (1.8)] (refining the Brauer–Siegel theorem), $|\mathrm{disc}(\mathrm{End}(E_2))|$ is polynomially bounded in terms of $[L(s) : L]$. Meanwhile Proposition 9.2 tells us that $N(s)$ is polynomially bounded in terms of $[L(s) : L]$. $\qquad\square$

Proof of Theorem 1.4. Let $\mathfrak{A} \to V$ be an abelian scheme satisfying the conditions of Theorem 1.4. This abelian scheme induces a morphism of varieties $q \colon V \to \mathcal{A}_2$, also defined over L. This morphism has finite fibres because V is a curve and $\mathfrak{A} \to V$ is non-isotrivial. Since $\mathrm{End}(\mathfrak{A}_{\overline{\eta}}) = \mathbb{Z}$, the image of V in $\mathcal{A}_2$ is not contained in a proper special subvariety (see section 2.F).

Hence Theorem 1.2 applies to the Zariski closure of the image of V in $\mathcal{A}_2$. Combining this with Proposition 9.3, we deduce that the image of V has finite intersection with Σ. We conclude because a point in V is an $E \times$ CM point if and only if its image lies in Σ. $\qquad\square$

If V is a curve satisfying the condition of Theorem 1.1 involving the Baily–Borel compactification, the following proposition allows us to construct an abelian scheme over a finite cover of V which satisfies the multiplicative reduction condition of Proposition 9.3, enabling us to deduce Theorem 1.1 from Proposition 9.3.

Proposition 9.4. *Let $V \subset \mathcal{A}_g$ be an irreducible algebraic curve such that the Zariski closure of V in the Baily–Borel compactification of $\mathcal{A}_g$ intersects the zero-dimensional stratum of the compactification. Then there exists a smooth projective curve $\tilde{V}'$, an open subset $\tilde{V} \subset \tilde{V}'$, a point $s_0 \in \tilde{V}' \setminus \tilde{V}$, a finite surjective morphism $q \colon \tilde{V} \to V$ and a semiabelian scheme $\mathfrak{A}' \to \tilde{V}'$ such that $\mathfrak{A}'_{|\tilde{V}}$ is an abelian scheme, the map to the coarse moduli space associated with $\mathfrak{A}'_{|\tilde{V}} \to \tilde{V}$ is the composition $\tilde{V} \xrightarrow{q} V \hookrightarrow \mathcal{A}_g$, and the fibre $\mathfrak{A}'_{s_0}$ is a torus.*
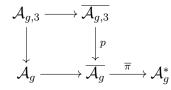
Proof. Let $\mathcal{A}_g^*$ be the Baily–Borel compactification of $\mathcal{A}_g$. Let V^* be the closure of V in $\mathcal{A}_g^*$. Recall that $\mathcal{A}_g^*$ has a stratification

$$\mathcal{A}_g^* = \mathcal{A}_g \sqcup \mathcal{A}_{g-1} \sqcup \cdots \sqcup \mathcal{A}_1 \sqcup \mathcal{A}_0.$$

By hypothesis, V^* intersects the stratum $\mathcal{A}_0$. This stratum is just a closed point, which we shall call s^*.

Let $\mathcal{A}_{g,3}$ denote the moduli space of principally polarised abelian varieties with full symplectic level-3 structure. Let $\overline{\mathcal{A}_g}$ and $\overline{\mathcal{A}_{g,3}}$ denote toroidal compactifications of $\mathcal{A}_g$ and $\mathcal{A}_{g,3}$ respectively, as defined in [FC90, Ch. IV, secs. 5 and 6]. We construct these toroidal compactifications with respect to some projective smooth admissible cone decomposition for $\mathbb{Z}^g$ (see [FC90, Ch. V, Definition 5.1] and remark (c) following that definition). Then $\overline{\mathcal{A}_g}$ is a proper algebraic stack [FC90, Ch. IV, Theorem 5.7] and $\overline{\mathcal{A}_{g,3}}$ is a projective algebraic variety [FC90, Ch. V, sec. 5]. Note that [FC90] constructs $\overline{\mathcal{A}_g}$ as a stack over $\mathbb{Z}$ and $\overline{\mathcal{A}_{g,3}}$ as a scheme over $\mathbb{Z}[\zeta_3, 1/3]$, but we require them only over $\mathbb{Q}$ and $\mathbb{Q}(\zeta_3)$ respectively.

We have the following commutative diagram of morphisms of stacks over $\mathbb{Q}(\zeta_3)$:

$$
\begin{array}{ccc}
\mathcal{A}_{g,3} & \longrightarrow & \overline{\mathcal{A}_{g,3}} \\
\downarrow & & \downarrow{\scriptstyle p} \\
\mathcal{A}_g & \longrightarrow & \overline{\mathcal{A}_g} \xrightarrow{\ \overline{\pi}\ } \mathcal{A}_g^*
\end{array}
$$

The composition $\overline{\pi} \circ p \colon \overline{\mathcal{A}_{g,3}} \to \mathcal{A}_g^*$ is a morphism of stacks between varieties, and consequently it is a morphism of varieties.

Let V_3 be an irreducible component of the preimage of V in $\mathcal{A}_{g,3}$ which surjects onto V. Let $\overline{V_3}$ be the Zariski closure of V_3 in $\overline{\mathcal{A}_{g,3}}$. This is a projective variety and hence $\overline{\pi} \circ p(\overline{V_3})$ is a closed subset of $\mathcal{A}_g^*$. Furthermore, $\overline{\pi} \circ p$ restricts to the natural map $\mathcal{A}_{g,3} \to \mathcal{A}_g$ and hence $\overline{\pi} \circ p(V_3) = V$. Therefore $s^* \in \overline{\pi} \circ p(\overline{V_3})$. Hence we can choose a closed point $\overline{s_3} \in \overline{V_3}$ such that $\overline{\pi} \circ p(\overline{s_3}) = s^*$.

By [FC90, Ch. IV, Theorem 5.7 (3)], there is a semiabelian scheme $\mathfrak{G}$ over the stack $\overline{\mathcal{A}_g}$ which extends the universal abelian scheme over $\mathcal{A}_g$ (considered as an algebraic stack). We can pull this back to a semiabelian scheme $\mathfrak{G}_3 \to \overline{\mathcal{A}_{g,3}}$. Define a stratification on $\overline{\mathcal{A}_g}$ by the dimensions of the abelian part of the fibres of $\mathfrak{G}$. According to [FC90, Ch. V, Theorem 2.3 (5)], $\overline{\pi}$ is compatible with the stratifications of $\overline{\mathcal{A}_g}$ and $\mathcal{A}_g^*$. It follows that $\mathfrak{G}_{3,\overline{s_3}}$ is a torus.

To finish the proof, let $\tilde{V}'$ be the normalisation of $\overline{V_3}$, let $\tilde{V}$ be the preimage of V_3 in $\tilde{V}'$, let s_0 be a preimage of $\overline{s_3}$ in $\tilde{V}'$ and let $\mathfrak{A}'$ be the pullback of $\mathfrak{G}_3$ to $\tilde{V}'$. $\square$

Proposition 9.5. *Let $V \subset \mathcal{A}_2$ be an algebraic curve satisfying the conditions of Theorem 1.1. Then Conjecture 1.3 holds for V.*

Proof. Let L be a number field over which V is defined. We can construct $\tilde{V}'$, $q \colon \tilde{V} \to V$, $s_0 \in \tilde{V}'$ and $\mathfrak{A}'$ as in Proposition 9.4. We can find a finite extension $\tilde{L}/L$ such that $\tilde{V}'$, $q \colon \tilde{V} \to V$, s_0 and $\mathfrak{A}' \to \tilde{V}'$ are all defined over $\tilde{L}$. The abelian scheme $\mathfrak{A}'_{|\tilde{V}} \to \tilde{V}$ and the point $s_0 \in \tilde{V}'(\overline{\mathbb{Q}})$ satisfy the conditions of Proposition 9.3. We can therefore complete the proof using Propositions 9.1 and 9.3. $\square$

Combining Theorem 1.2 and Proposition 9.5 proves Theorem 1.1.

References

[And89] Y. André, "G-functions and geometry", Aspects of Mathematics, E13, Friedr. Vieweg & Sohn, Braunschweig, 1989.

[BHC62] A. Borel and Harish-Chandra, *Arithmetic subgroups of algebraic groups*, Ann. of Math. (2) **75** (1962), 485–535.

[BLR90] S. Bosch, W. Lütkebohmert, and M. Raynaud, "Néron models", Ergeb. Math. Grenzgeb. (3), vol. 21, Springer-Verlag, Berlin, 1990.

[Bor84] M. V. Borovoĭ, *Langlands' conjecture concerning conjugation of connected Shimura varieties*, Selecta Math. Soviet. **3** (1983/84), no. 1, 3–39.

[BS91] M. Burger and P. Sarnak, *Ramanujan duals. II*, Invent. Math. **106** (1991), no. 1, 1–11.

[Clo03] L. Clozel, *Démonstration de la conjecture τ*, Invent. Math. **151** (2003), no. 2, 297–328.

[Del79] P. Deligne, *Variétés de Shimura: interprétation modulaire, et techniques de construction de modèles canoniques*, In: "Automorphic forms, representations and *L*-functions (Part 2)", Proc. Sympos. Pure Math., XXXIII, AMS, Providence, R.I., 1979, 247–289.

[DO16] C. Daw and M. Orr, *Heights of pre-special points of Shimura varieties*, Math. Ann. **365** (2016), no. 3-4, 1305–1357.

[DO] C. Daw and M. Orr, *Quantitative reduction theory and unlikely intersections*, preprint, available at arXiv:1911.05618v2.

[DR18] C. Daw and J. Ren, *Applications of the hyperbolic Ax–Schanuel conjecture*, Compos. Math. **154** (2018), 1843–1888.

[DS05] F. Diamond and J. Shurman, "A first course in modular forms", Grad. Texts in Math., vol. 228, Springer-Verlag, New York, 2005.

[Fal83] G. Faltings, *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*, Invent. Math. **73** (1983), no. 3, 349–366.

[FC90] G. Faltings and C.-L. Chai, "Degeneration of abelian varieties", Ergeb. Math. Grenzgeb. (3), vol. 22, Springer-Verlag, Berlin, 1990, with an appendix by D. Mumford.

[HP12] P. Habegger and J. Pila, *Some unlikely intersections beyond André–Oort*, Compos. Math. **148** (2012), no. 1, 1–27.

[HP16] P. Habegger and J. Pila, *O-minimality and certain atypical intersections*, Ann. Sci. Éc. Norm. Supér. (4) **49** (2016), no. 4, 813–858.

[Huy16] D. Huybrechts, "Lectures on K3 surfaces", Cambridge Stud. Adv. Math., vol. 158, Cambridge University Press, Cambridge, 2016.

[Jar14] F. Jarvis, "Algebraic number theory", Springer Undergrad. Math. Ser., Springer, 2014.

[Kow] E. Kowalski, "An introduction to expander graphs", Online notes, available at `https://people.math.ethz.ch/~kowalski/`.

[Len87] H. W. Lenstra, Jr., *Factoring integers with elliptic curves*, Ann. of Math. (2) **126** (1987), no. 3, 649–673.

[LM16] H. Li and G. A. Margulis, *Effective estimates on integral quadratic forms: Masser's conjecture, generators of orthogonal groups, and bounds in reduction theory*, Geom. Funct. Anal. **26** (2016), no. 3, 874–908.

[LZ89] A. Lubotzky and R. J. Zimmer, *Variants of Kazhdan's property for subgroups of semisimple groups*, Israel J. Math. **66** (1989), no. 1-3, 289–299.

[LZ03] A. Lubotzky and A. Żuk, "On property (τ)", Preliminary version, available at `http://www.ma.huji.ac.il/~alexlub/`, 2003.

[Mas] D.W. Masser, *Semi–effective André–Oort for real multiplication curves in the Siegel threefold*, Preliminary version sent via private communication.

[Mil83] J. S. Milne, *The action of an automorphism of* **C** *on a Shimura variety and its special points*, In: "Arithmetic and geometry, Vol. I", M. Artin and J. Tate (eds.), Progr. Math., vol. 35, Birkhäuser Boston, Boston, MA, 1983, 239–265.

[Mil05] J. S. Milne, *Introduction to Shimura varieties*, In: "Harmonic analysis, the trace formula, and Shimura varieties", J. Arthur, D. Ellwood and R. Kottwitz (eds.), Clay Math. Proc., vol. 4, Amer. Math. Soc., Providence, RI, 2005, 265–378.

[MS82] J. S. Milne and K. Y. Shih, *Conjugates of Shimura varieties*, In: "Hodge cycles, motives, and Shimura varieties", Lecture Notes in Math., vol. 900, Springer-Verlag, 1982, 280–356.

[Mum74] D. Mumford, "Abelian varieties", second ed., Oxford University Press, 1974.

[MW93] D. Masser and G. Wüstholz, *Isogeny estimates for abelian varieties, and finiteness theorems*, Ann. of Math. (2) **137** (1993), no. 3, 459–472.

[MZ99] B. J. J. Moonen and Yu. G. Zarhin, *Hodge classes on abelian varieties of low dimension*, Math. Ann. **315** (1999), no. 4, 711–733.

[Orr18] M. Orr, *Height bounds and the Siegel property*, Algebra Number Theory **12** (2018), no. 2, 455–478.

[Orr19] M. Orr, *Galois conjugates of special points and special subvarieties in Shimura varieties*, J. Inst. Math. Jussieu, published online 2019-10-30.

[Orr] M. Orr, *Unlikely intersections with Hecke translates of a special subvariety*, to appear in J. Eur. Math. Soc. (JEMS), available at arXiv:1710.04092v2.

[PT13] J. Pila and J. Tsimerman, *The André–Oort conjecture for the moduli space of abelian surfaces*, Compos. Math. **149** (2013), no. 2, 204–216.

[PT14] J. Pila and J. Tsimerman, *Ax–Lindemann for $\mathcal{A}_g$*, Ann. of Math. (2) **179** (2014), no. 2, 659–681.

[Tsi12] J. Tsimerman, *Brauer–Siegel for arithmetic tori and lower bounds for Galois orbits of special points*, J. Amer. Math. Soc. **25** (2012), no. 4, 1091–1117.

[Tsi18] J. Tsimerman, *The André–Oort conjecture for $\mathcal{A}_g$*, Ann. of Math. (2) **187** (2018), no. 2, 379–390.

[Ull07] E. Ullmo, *Equidistribution de sous-variétés spéciales II*, J. Reine Angew. Math. **606** (2007), 193–216.

[UY15] E. Ullmo and A. Yafaev, *Nombre de classes des tores de multiplication complexe et bornes inférieures pour les orbites galoisiennes de points spéciaux*, Bull. Soc. Math. France **143** (2015), no. 1, 197–228.

[Vas68] L. N. Vaserstein, *On groups possessing property T*, Funktsional. Anal. i Prilozhen. **2** (1968), no. 2, 86 (English translation: Funct. Anal. Appl. **2** (1968), no. 2, 174).

Daw: Department of Mathematics and Statistics, University of Reading, Whiteknights, PO Box 217, Reading, Berkshire RG6 6AH, United Kingdom
Email address: `chris.daw@reading.ac.uk`

Orr: Mathematics Institute, Zeeman Building, University of Warwick, Coventry CV4 7AL, United Kingdom
Email address: `martin.orr@warwick.ac.uk`