J. Eur. Math. Soc. 24, 3161–3182 (2022) DOI 10.4171/JEMS/1161

Jonathan Pila · Jacob Tsimerman

Independence of CM points in elliptic curves

Received July 5, 2019

Abstract. We prove a result which describes, for each $n \geq 1$, all linear dependencies among n images in a given elliptic curve of special points in a given modular or Shimura curve under a given parameterization (or correspondence). Our result unifies and improves in certain aspects previous work of Rosen–Silverman–Kühne and Buium–Poonen.

Keywords. Zilber–Pink, Heegner point, unlikely intersection

1. Introduction and main results

Let Y be a modular (or Shimura) curve, E an elliptic curve over $\mathbb{C}$, and $V \subset Y \times E$ an irreducible correspondence (i.e. an irreducible closed subvariety of $Y \times E$ of dimension 1 dominant to both factors). If $(s, x) \in V$ we will call x a V*-image* of s. We prove a result describing, for each $n \geq 1$, all linear dependencies in E among the V-images of n special points (see below) in Y. An example of particular interest is when V is the graph of a modular parameterization $\phi : Y \to E$ and then the V-images of special points are known as *CM points* or *Heegner points* (though the latter term is usually taken to involve some further assumptions). A number of results in the literature establish linear independence of CM points under suitable hypotheses. After framing our result we compare it with previous results.

The *special subvarieties* in E^n are the cosets of abelian subvarieties by torsion points ("torsion cosets"); the *special subvarieties* of Y^n when Y is the modular curve $Y(1)$ are described e.g. in [13, 3.2]; they are fibre products of special points and modular curves. In general, see [20, Definition 2.5]. The special subvarieties of Y^n are described in §2. *Special points* in either ambient are special subvarieties of dimension zero: torsion points in E^n, and in $Y(1)^n$, tuples of singular moduli. We identify varieties with their sets of complex points and subvarieties are assumed to be relatively closed.

Definition. With notation as above, and $n \geq 1$, let π_{Y^n}, π_{E^n} be the projections of $Y^n \times E^n$ onto the first and second factors, respectively.

Jonathan Pila: Mathematical Institute, University of Oxford, Oxford, UK; pila@maths.ox.ac.uk
Jacob Tsimerman: Department of Mathematics, University of Toronto, Toronto, Canada; jacobt@math.toronto.edu

Mathematics Subject Classification (2020): Primary 11G18, 11G05; Secondary 03C64, 14G35

(i) A *distinguished component* in V^n is a (geometrically) irreducible component $W \subset V^n \cap (S \times B)$, where $S \subset Y^n$ and $B \subset E^n$ are special subvarieties, such that $\pi_{Y^n}(W) = S$ and $\pi_{E^n}(W) \subset B$.

(ii) A distinguished component W in V^n is called *dependent* if the special subvariety B may be taken so that the inclusion $B \subsetneq E^n$ is strict.

(iii) A distinguished component W in V^n is called *exemplary* if, setting B to be the smallest special subvariety of E^n with $\pi_{E^n}(W) \subset B$, there is no distinguished component W' strictly larger than W with $\pi_{E^n}(W') \subset B$.

Observe that when V is the graph of a parameterization $\phi : Y \to E$, a distinguished component is simply the graph of the restriction of ϕ to a special subvariety $S \subset Y^n$. Note that an exemplary component need not be dependent, but the unique non-dependent exemplary component is V^n itself, being a component of $V^n \cap (Y^n \times E^n)$.

Let $x_1, \dots, x_n \in E$ be V-images of special points $s_1, \dots, s_n \in Y$. Write $s = (s_1, \dots, s_n) \in Y^n$, $x = (x_1, \dots, x_n) \in E^n$. If $(s, x) \in W$ for some dependent distinguished component in V^n, then the points $x_1, \dots, x_n \in E$ are linearly dependent in E. Note that, for us, *linear dependence in* E is always taken to be over $\mathrm{End}(E)$. We have $\mathrm{End}(E) = \mathbb{Z}$ unless E has CM (complex multiplication), in which case $\mathrm{End}(E)$ is an order in an imaginary quadratic field. Conversely, if $x_1, \dots, x_n$ are linearly dependent in E then (s, x) is contained in some dependent exemplary component.

The following theorem thus gives a description of every linear dependence among V-images of n special points.

Theorem 1.1. *Let $V \subset Y \times E$ be as above with Y a modular curve or a Shimura curve and $n \geq 1$. Then there are only finitely many exemplary components in V^n.*

Note. We have assumed that V is dominant to both factors. However, only the dominance to Y is needed in Theorem 1.1. Suppose $V = Y \times \{x\}$. Then for each n, one has the unique exemplary component $(Y \times \{x\})^n$, which is dependent if and only if x is torsion.

Example. It is well known that $X_0(11)$ has the structure of an elliptic curve, so we may set $E = \mathrm{Pic}^0(X_0(11))$. Consider the Atkin–Lehner involution w. Now on E, w is a non-trivial automorphism, so its graph must be an abelian subvariety. Set $\phi : X_0(11) \to E$ to be the identification taking $(\infty) \to 0$. Then $\phi(w(\infty)) = \phi(0)$ is a torsion point, and thus if we set $S \subset X_0(11)^2$ to be the graph of w, then S is a special curve whose ϕ-graph is exemplary.

A number of results in the literature assert linear independence properties of the V-images of CM points. The fact that only finitely many V-images of special points can be torsion was proved in [21, Theorem 1.5] for modular parameterizations and Heegner points (generalized to certain Shimura curve parameterizations in [14, Theorem 1]) and is equivalent to the assertion of Theorem 1.1 for $n = 1$. This also follows from the stronger results in [5], in particular Theorem 1.1 there, and was reproved as a "special point problem" within the Zilber–Pink conjecture in [25, Theorem 1.3].

We deduce some consequences of Theorem 1.1 and compare with some further results in the literature. It is convenient for the statement to order the (countable set of) special, non-fibral curves of Y^2. When Y is the modular curve, we let $X_N \subset Y \times Y$ for $N \geq 1$ be the classical modular curve of level N, i.e. the locus of points (s_1, s_2) such that there is a cyclic isogeny of degree N between the corresponding elliptic curves. When Y is not the modular curve, we fix $(X_N)_{N \in \mathbb{N}}$ to be any ordering of our choosing.

The *discriminant* $\Delta(s)$ of a special point $s \in Y(1)$ is the discriminant of the quadratic order which is the endomorphism ring of the corresponding CM elliptic curve; in general see §5, above 5.1. The size of the discriminant measures the "complexity" of a special point.

Definition. Let D be a positive integer. A set $\{s_1, \ldots, s_n\}$ of special points in Y is called *D-independent* if, for each i, $|\Delta(s_i)| > D$ and, for $i \neq j$, there is no relation $(s_i, s_j) \in X_N$ with $N \leq D$.

Corollary 1.2. *For $n \geq 1$ there exists a positive integer $D = D(Y, E, V, n)$ such that if $\{s_1, \ldots, s_n\}$ is D-independent then any V-images $x_1, \ldots, x_n$ of $s_1, \ldots, s_n$ are linearly independent in E.*

Note. Here and below we do require that V is dominant to E, so that any dependent component must arise from a proper special subvariety of Y^n.

Proof of Corollary 1.2. For $s = (s_1, \ldots, s_n)$ to have a V-image which is dependent requires s to lie in one of finitely many proper special subvarieties $S_1, \ldots, S_k \subset Y^n$. We claim the maximal, non-fibral proper special subvarieties are all pulled back from non-fibral special curves in Y^2 for some pair of coordinates. To see this, note that any proper semisimple subgroup of SL_2^n – which is surjective on all coordinate projections – projects to the graph of an isomorphism on some two coordinates. Since special subvarieties are orbits of semisimple subgroups, the claim follows. Thus, for each i, $s \in S_i$ requires either that some coordinate is equal to a fixed special point, or that $(s_i, s_j) \in X_N$ for some $i \neq j$ and fixed N. This is not possible if s is D-independent for sufficiently large D. ∎

Corollary 1.2 improves a result of Kühne [15, Theorem 2], which in turn improves, in some respects, Rosen–Silverman's [32, Theorem 1]. The results differ in detail and so it is not straightforward to compare them. With respect to Kühne's result, ours implies independence even for CM points corresponding to orders in the same CM field, if the orders are sufficiently far apart (i.e. if the corresponding singular moduli are D-independent up to suitable D; the previous results required the CM fields of the s_i to be distinct). Note however that Kühne's result is effective, whereas ours is not. With respect to [32], Kühne's result ameliorates the condition on the class number in [32, Theorem 1], which was insufficient to imply finiteness. But note that the constant $C(E)$ in [32, Theorem 1] is independent of n, a feature not recovered in subsequent results. The results in [15,32] also exclude CM elliptic curves E (though see [33]), and all these results restrict to modular parameterizations of $E/\mathbb{Q}$ (and [32] restricts to CM by maximal orders).

Let Σ denote the set of V-images in E of special points of Y.

Corollary 1.3. *For $n \geq 1$ there exists a positive integer $N = N(Y, E, V, n)$ such that if $x_1, \dots, x_N \in \Sigma$ are distinct then there is a linearly independent subset of $\{x_i\}$ of size at least n.*

Proof. Given n we can find N such that any set of N distinct V-images of special points contains a subset of size n for which the corresponding special points are $D(Y, E, V, n)$-independent. (And N is effective given D.) ∎

Corollary 1.4. *Let Γ be a finitely generated subgroup of E of rank r. Then $|\Gamma \cap \Sigma| \leq N(Y, E, V, r+1)$.* ∎

This re-proves Buium–Poonen's [5, Theorem 1.1] (and generalizes to correspondences their result (Theorem 2.5) for maps from Shimura curves to elliptic curves) and in a uniform way: the size of the intersection is bounded depending only on the rank of Γ. However, we cannot recover their Bogomolov-type result.

In §2 we show that Theorem 1.1 is a consequence of the Zilber–Pink conjecture (ZP). The framing of ZP in terms of "optimal subvarieties" (as in [13]; see §2) suggests the formulation of Theorem 1.1. We are not able however to prove the full ZP statement for V^n; see the comments before Proposition 2.1.

Our proof of Theorem 1.1 goes via point-counting on definable sets in o-minimal structures, and utilizes a suitable Ax–Schanuel theorem, as in various earlier works dealing with special cases of ZP, and in this respect follows in particular the approach in [27] in studying "CM-points" for the multiplicative group. As there, various issues arise from the fact that we cannot prove the full Zilber–Pink statement for V^n. But unlike in [27], where we showed that no positive-dimensional dependent distinguished components exist, we must here deal with this possibility, which complicates the point-counting and the application of Ax–Schanuel, in view of our inability to affirm the full ZP. We must show that we are able to restrict throughout to atypical intersections of a specific form.

In effect, we must prove a very strong result of André–Oort type: each proper special subvariety of E^n has a preimage in Y^n. This gives a countably infinite collection of subvarieties of Y^n which is not contained in any algebraic family. We must show that there are only finitely many special subvarieties of Y^n which are contained and maximal in any one of this countably infinite collection.

In the modular case we show that our results can be extended to include the Hecke orbits of a finite number of points in addition to special points. The *Hecke orbit* of $u \in Y$ is $\{v \in Y : \exists N : (u, v) \in X_N\}$.

Definition. Let Y be a modular curve and $U \subset Y$.

(i) A *U-special point* of Y is a point which is either special or in the Hecke orbit of some $u \in U$.

(ii) A *U-special point* in Y^n is an n-tuple of U-special points in Y.

(iii) A *U-special subvariety* of Y^n is a weakly special subvariety which contains a U-special point.

Now we consider again an irreducible correspondence $V \subset Y \times E$.

Definition. Let notation be as above.

(i) A *distinguished U-component* in V^n is a component $W \subset V^n \cap (S \times B)$, where $S \subset Y^n$ is U-special, $B \subset E^n$ is special, $\pi_{Y^n}(W) = S$, and $\pi_{E^n}(W) \subset B$.

(ii) A distinguished U-component W in V^n is *dependent* if B (may be taken such that it) is a proper special subvariety.

(iii) A distinguished U-component W is *exemplary* if, setting B to be the smallest special subvariety of E^n with $\pi_{E^n}(W) \subset B$, there is no distinguished U-component W' strictly larger than W with $\pi_{E^n}(W') \subset B$.

Theorem 1.5. *Given $V \subset Y \times E$ as above with Y a modular curve, $U \subset Y$ finite, and $n \geq 1$, there are only finitely many exemplary U-components in V^n.*

One may deduce corollaries analogous to 1.2–1.4 above. The last recovers a result of Baldi [1, Theorem 1.3], (obtained via equidistribution), which is also a special case of results of Dill [9, 10], affirming a conjecture of Buium–Poonen [6, Conjecture 1.7]; see the discussion in [1]. Baldi obtains a stronger “Bogomolov”-type result, which we do not. These results are in the circle of the André–Pink conjecture (see [29] and further references in [1]), though Theorem 1.5 is rather an “unlikely intersection” result in such contexts. Of course it too is subsumed under the general Zilber–Pink conjecture.

With existing arithmetic estimates, Theorem 1.5 and its corollaries should generalize to Shimura curves, with a suitable notion of Hecke orbit.[1]

The structure of the paper is as follows. In §2 we define Shimura curves for our purposes and gather facts we need about them. The Zilber–Pink setting is recalled in §3. The Ax–Schanuel statement and refinements we need are given in §4. Some arithmetic estimates are collected in §5. Theorems 1.1 and 1.5 are proved in §6 when everything is defined over a number field, and extended to $\mathbb{C}$ in §7. In this paper, *definable* will mean “definable in the o-minimal structure $\mathbb{R}_{\text{an, exp}}$”; for background on o-minimality and on $\mathbb{R}_{\text{an, exp}}$ see [24].

2. Shimura curves

2.1. Definitions

Definition. A *Shimura curve* simply means a connected Shimura variety of dimension one, which is of *abelian type*.[2]

We define $\mathcal{A}_g$ to be the moduli space of principally polarized abelian varieties of dimension g. Recall that a Shimura variety S is *Hodge type* if S admits an injective

[1] There is an issue with abelian varieties that one could consider isogenies not necessarily respecting the polarization, which complicates matters.

[2] These include all quaternionic Shimura curves, as well as those studied by Deligne [8, Section 6]. The authors do not know whether all one-dimensional Shimura varieties are of abelian type.

homomorphism into $\mathcal{A}_g$. A Shimura variety S is said to be of *abelian type* if it admits an isogeny to a Shimura variety of Hodge type.

We let Y denote an arbitrary Shimura curve, and state and prove some basic facts. These are probably known to experts but we collect them here for the reader's convenience.

Recall first that Y is uniformized by the upper half-plane $\mathbb{H}$, and can be written as $\Gamma\backslash\mathbb{H}$ for some discrete arithmetic subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{R})$.

By fixing an isogeny to a Shimura variety of Hodge type, we may therefore associate to every special point on Y a CM principally polarized abelian variety.

2.2. *Weakly special subvarieties*

In this section we determine the structure of weakly special subvarieties of Y^n. In particular, we prove that there are only non-trivial relations on two variables at a time. As in the introduction, let $(X_N)_{N\in\mathbb{N}}$ be any ordering of the special, non-fibral curves in Y^2.

Lemma 2.1. *Let $V \subset Y^n$ be a weakly special variety. Then there is a partition of $\{1,\ldots,n\}$ into disjoint subsets $K, I_1,\ldots,I_n$ (in which K only is permitted to be empty) such that:*

- *For any $k \in K$, the projection $\pi_k : V \to Y^{\{k\}}$ is constant.*
- *For any $a \neq b$ belonging to the same I_j, the image of the projection $\pi_{a,b} : V \to Y^{\{a,b\}}$ is special.*

Proof. First, note that a non-fibral weakly special curve in Y^2 must be special. This follows from the fact that the corresponding semisimple group is a conjugate of the diagonal SL_2 in SL_2^2 and is therefore its own normalizer.

We can assume without loss of generality that none of the coordinates are constant on Y, by projecting them away. Now set $d = \dim V$, and let $\{n_1,\ldots,n_d\}$ be indices such that $z_{n_1},\ldots,z_{n_d}$ are algebraically independent over V. We will prove that for any other coordinate z_r, there exists an index i such that z_r is algebraic over z_{n_i}.

To see this, let $J \subset \{n_1,\ldots,n_d\}$ be a minimal set of indices such that z_r is algebraic over $\mathbb{C}(z_j)_{j\in J}$. Assume for the sake of contradiction that $|J| \geq 2$. Since fibres of weakly special varieties are weakly special, by fibring over $z_j = c$ we obtain a family of weakly special varieties V_c. Moreover, for each $c \in \mathbb{C}$ it follows that z_r is algebraic over z_{j_c} for some $j_c \in J$. This implies that $\pi_{r,j_c}(V_c)$ is a finite union of non-fibral weakly special curves. Now we may choose an index $j \in J$ and a special curve X_N such that for infinitely many values of $c \in \mathbb{C}$ we have $X_N \subset \pi_{r,j}(V_c)$. But now since V is irreducible, it follows that $V = \pi_{r,j}^{-1}(X_N)$, contradicting the minimality of J. ∎

A *special point* of Y is one for which the corresponding elliptic curve or abelian variety is CM, and a *special subvariety* is a weakly special subvariety which contains a special point; equivalently, in the above description, any coordinate which is constant is special.

Finally, we shall make use of the following definition to group the weakly specials into natural families.

Definition. We define the *slope* of a weakly special variety $V \subset Y^n$ to be the partition $K, I_1, \ldots, I_n$ given by Lemma 2.1, as well as the images $\pi_{a,b}(V)$ for each pair $a \neq b$ belonging to the same I_j.

3. The Zilber–Pink setting

We place Theorem 1.1 in the context of the Zilber–Pink conjecture (ZP) proposed independently, in slightly different formulations, by Zilber [36], Bombieri–Masser–Zannier [3], and Pink [30].

This concerns a (*connected*) *mixed Shimura variety* M and its collection $\mathcal{S}$ of *special subvarieties*. One also has the larger collection of *weakly special subvarieties*. For definitions see e.g. Pink [29, Definitions 2.1, 4.1] and Gao [11, Definition 3.9]. Let $Z \subset M$ be a subvariety. For $S \in \mathcal{S}$, a component $A \subset Z \cap S$ is *atypical* if

$$\dim A > \dim Z + \dim S - \dim M.$$

(The intersection is called *unlikely* if $\dim Z + \dim S - \dim M < 0$.) ZP predicts a description in finite terms of all "atypical" intersections of Z with special subvarieties $S \in \mathcal{S}$.

For a subvariety $Z \subset M$ we let $\langle Z \rangle$ denote the smallest special subvariety of M containing Z, and by $\langle Z \rangle_{\mathrm{ws}}$ the smallest weakly special one.

We define the *defect* $\delta(Z)$ of Z and the *weakly special defect* $\delta_{\mathrm{ws}}(Z)$ by

$$\delta(Z) = \dim \langle Z \rangle - \dim Z, \quad \delta_{\mathrm{ws}}(Z) = \dim \langle Z \rangle_{\mathrm{ws}} - \dim Z.$$

Definition. Let $Z \subset M$.

(i) A subvariety $A \subset Z$ is called *optimal* if it is maximal for its defect as a subvariety of Z. That is, if $A \subset B \subset Z$ and $\delta(B) \leq \delta(A)$ then $B = A$.

(ii) A subvariety $A \subset Z$ is called *geodesic optimal* if it is maximal for its weakly special defect as a subvariety of Z.

The following is formally equivalent to the strongest form of ZP, namely the analogue for a mixed Shimura variety of the conjectures of Zilber and Bombieri–Masser–Zannier (for semiabelian varieties and $\mathbb{G}_{\mathrm{m}}$), as shown in [13]. (The notion here called "geodesic optimal" was earlier introduced as "cd-maximal" in a different context in [31] in the setting of $\mathbb{G}_{\mathrm{m}}$.)

Conjecture 3.1 (ZP). *Let $Z \subset M$. Then Z has only finitely many optimal subvarieties.*

The ambient variety $Y^n \times E^n$ is an example of a *weakly special subvariety* of a *mixed Shimura variety* (it is *special* precisely if E has CM). Namely, let

$$\mathcal{E} \to Y$$

be the universal family over Y (of elliptic curves if Y is a modular curve, or of abelian surfaces if Y is a Shimura curve). Then $\mathcal{E}$ is a mixed Shimura variety (see e.g. [11]), in which Y can be identified with the zero-section. If E is isomorphic to the fibre over

$s \in Y$ then it may be identified with this fibre, which is weakly special. Correspondingly, $Y^n \times E^n$ may be identified with a weakly special subvariety of $\mathcal{E}^n \times \mathcal{E}^n$.

It is well-known (see e.g. Pink [30]) that ZP implies a similar statement for its weakly special subvarieties, whose "special subvarieties" are simply the intersections of it with special subvarieties of the ambient mixed Shimura variety. There are corresponding notions of smallest special and weakly special subvariety containing a given subvariety, defect and weakly special defect, and ZP can be expressed in terms of the corresponding notion of "optimal" as above; in what follows, the notation $\langle\cdot\rangle$ and defects will always be with respect to the ambient variety $Y^n \times E^n$.

Proposition 3.1. *Let S be a pure Shimura variety of abelian type and T an abelian variety. Then the (weakly) special subvarieties of $S \times T$ are precisely the products $U \times B$ of (weakly) special $U \subset Y^n$, $B \subset T$.*

Proof. There is a Shimura morphism $S \to S'$ to a Hodge-type pure Shimura variety $S' \subset \mathcal{A}_g$. Say $\dim T = h$. We may identify $S' \times T$ with a weakly special subvariety of the product of mixed Siegel modular varieties, $\mathcal{X}_g \times \mathcal{X}_h \to \mathcal{A}_g \times \mathcal{A}_h$. Let π be the projection under which T projects to a point in $\mathcal{A}_h$. By [11, Proposition 3.7], a (weakly) special subvariety of $\mathcal{X}_g$ is (up to a finite cover of $\pi(S')$) a translate of an abelian subscheme of $\pi^{-1}(\pi(S')) \to \pi(S')$ by a torsion section and then by a constant section over a weakly special subvariety $\pi(S') \subset \mathcal{A}_g$. It follows that (weakly) special subvarieties of $S \times T$ are products of (weakly) special subvarieties of the factors. The converse always holds. ■

Presumably the same holds for any product $P \times T$ of a pure Shimura variety and an abelian variety.

It follows then that, for $Z \subset Y^n \times E^n$,

$$\langle Z\rangle = \langle \pi_{Y^n}(Z)\rangle_{Y^n} \times \langle \pi_{E^n}(Z)\rangle_{E^n}$$

and likewise for $\langle Z\rangle_{\mathrm{ws}}$.

Given $V \subset Y \times E$, we consider ZP for $V^n \subset Y^n \times E^n$. If $x \in E^n$ is a V-image of a special point $s \in Y^n$ and x is dependent then $x \in B$ for some proper special subvariety of E^n. Then $(s, x) \in V^n \cap (\{s\} \times B)$, and since $\dim(\{s\} \times B) + \dim V^n < 2n$ this shows that any dependent image of a special point is an "unlikely" or "atypical" intersection in the sense of the Zilber–Pink conjecture.

The following shows that exemplary components are optimal subvarieties of V^n, and hence that Theorem 1.1 is a consequence of ZP. However, we are not able to prove ZP for V^n (once $n \geq 3$); the analogous situation when E is replaced by the multiplicative group $\mathbb{G}_{\mathrm{m}}$ is discussed in detail in [27].

Proposition 3.2. *An exemplary component in V^n is an optimal subvariety of V^n.*

Proof. Let $W \subset V^n \cap (S \times B)$ be an exemplary component with $\pi_{Y^n}(W) = S$ and $B = \langle \pi_{E^n}(W)\rangle$. Then $\dim W = \dim S$ and the smallest special subvariety of $Y^n \times E^n$ containing W is $S \times B$. Hence the defect of W is

$$\delta(W) = \dim\langle W\rangle - \dim W = \dim S + \dim B - \dim W = \dim B.$$

If W were not optimal, it would be contained in some larger subvariety $W' \subset V^n$ of the same, or lower defect. Write

$$\langle W' \rangle = S' \times B'.$$

Then $B \subset B'$ and $\dim W' \le \dim S'$ and

$$\delta(W') = \dim \langle W' \rangle - \dim W' = \dim S' + \dim B' - \dim W'.$$

If $\delta(W') \le \delta(W)$ we must have $B' = B$ and $\dim W' = \dim S'$, which would mean that W' is a distinguished component in V^n on S', containing W, projecting into B. But by the maximality of W we have $W' = W$. ■

We will need the "weak" analogue of the above. A *weakly distinguished component* in V^n is a component $W \subset V^n \cap (S \times B)$ where S, B are weakly special subvarieties. It is *weakly exemplary* if, taking $B = \langle \pi_{E^n}(W) \rangle_{\mathrm{ws}}$, there is no weakly distinguished component W' strictly larger than W with $\pi_{E^n}(W') \subset B$.

Proposition 3.3. *A weakly exemplary component in V^n is a geodesic optimal subvariety of V^n.*

Proof. The same as in Proposition 3.2. ■

The Ax–Schanuel theorem only detects weakly special subvarieties, and we thus need to show (as has already been shown in several other settings, including for all pure Shimura varieties by Daw–Ren [7]) that optimal subvarieties are geodesic optimal. For this we establish the "defect condition".

Definition. A weakly special subvariety X of a mixed Shimura variety has the *defect condition* if, for $A \subset B \subset X$, we have

$$\delta(B) - \delta_{\mathrm{ws}}(B) \le \delta(A) - \delta_{\mathrm{ws}}(A),$$

the defects being with respect to the special and weakly special subvarieties of X.

Proposition 3.4. *Let S be a pure Shimura variety of abelian type and T an abelian variety. Then $S \times T$ has the defect condition.*

Proof. For an abelian variety (as well as for $\mathbb{G}_{\mathrm{m}}^n$ and products of modular curves) the defect condition is established in [13, Proposition 4.3], and for a general pure Shimura variety in [7, 4.4]. By Proposition 3.1, the (weakly) special subvarieties of $S \times T$ are products of (weakly) special subvarieties of the factors. Thus we have

$$\langle A \rangle = \langle \pi_S(A) \rangle_S \times \langle \pi_T(A) \rangle_T$$

so that

$$\delta(A) - \delta_{\mathrm{ws}}(A) = \delta(\pi_S(A)) - \delta_{\mathrm{ws}}(\pi_S(A)) + \delta(\pi_T(A)) - \delta_{\mathrm{ws}}(\pi_T(A)),$$

and likewise for B, and the defect condition for $S \times T$ follows from the defect conditions in S and T by addition. ■

It is conjectured in [13] that the defect condition holds in all mixed Shimura varieties.

Proposition 3.5. *An optimal subvariety is geodesic optimal.*

Proof. This follows formally once one has the defect condition, as in [13]. ∎

4. An Ax–Schanuel result

The *Ax–Schanuel property* for the uniformization map

$$u_M : D \to M$$

realizing a mixed Shimura variety M as a quotient of a suitable Hermitian symmetric domain D by a discrete arithmetic group Γ is a functional transcendence statement for u_M analogous to the classical Ax–Schanuel theorem for the exponential function $\exp : \mathbb{C} \to \mathbb{C}^\times$. For discussion and proof of such results see [11,19]. Such a result implies a corresponding statement for each weakly special subvariety $X \subset M$, uniformized by an irreducible component of $u_M^{-1}(X)$.

The Ax–Schanuel result we need is for the uniformization

$$u : \mathbb{H} \times \mathbb{C} \to Y \times E$$

and its cartesian powers. We will use the same notation

$$u : \mathbb{H}^n \times \mathbb{C}^n \to Y^n \times E^n$$

for cartesian powers of u.

We will (as usual in ZP applications) use only the "two-sorted" form, which we now state, after noting the following convention. Strictly speaking, $\mathbb{H}^n$ has no algebraic subvarieties. By an *algebraic subvariety* of U, where $U \subset \mathbb{H}^n \times \mathbb{C}^n$ is a weakly special subvariety, we will mean an irreducible analytic component of the intersection of U with an algebraic subvariety (in the usual sense) of the ambient $\mathbb{C}^n \times \mathbb{C}^n$.

Theorem 4.1. *Let U' be a weakly special subvariety of $\mathbb{H}^n \times \mathbb{C}^n$ with image $u(U') = X'$ a weakly special subvariety of $Y^n \times E^n$. Let $Z \subset X'$, $A \subset U'$ be algebraic varieties, and Ω an irreducible analytic component of $A \cap u^{-1}(Z)$. Then*

$$\dim \Omega = \dim Z + \dim A - \dim X'$$

unless Ω is contained in a proper weakly special subvariety of U'.

Proof. The Ax–Schanuel statement is easily seen to be invariant under isogenies of the Shimura varieties, and the abelian varieties. We therefore assume throughout the proof that Y is a special subvariety of $\mathcal{A}_g$.

Since the statement is about the uniformization corresponding to a weakly special subvariety of $\mathcal{E}^{2n}$, the result follows from the Ax–Schanuel statement for the uniformization

$$\mathbb{H}^{2n} \times \mathbb{C}^{2n} \to \mathcal{E}^{2n}$$

and since Y^{2n}, the "pure" Shimura variety underlying $\mathcal{E}^{2n}$, is a special subvariety of $\mathcal{A}_g$, the Siegel modular variety of principally polarized abelian varieties, when $g \geq 2n$ the required Ax–Schanuel follows from the corresponding statement for the universal family $\mathcal{X}_g$ of abelian varieties over $\mathcal{A}_g$, namely the Ax–Schanuel theorem for the uniformization

$$\mathbb{H}_g \times \mathbb{C}^g \to \mathcal{X}_g.$$

This theorem is due to Gao [11, Theorem 1.1], extending, for $\mathcal{A}_g$, the result for a general pure Shimura variety in [19, Theorem 1.1]. ∎

As in [7, 13], this can be reformulated in terms of a suitable notion of "optimality", for which we adopt the terminology used by Daw–Ren [7, §§5.7–5.9], to distinguish it from "optimality" as above in §2.

Definition. Let $Z \subset Y^n \times E^n$ be a subvariety.

(i) An *intersection component* of $u^{-1}(Z)$ is an irreducible analytic component of the intersection of $u^{-1}(Z)$ with an algebraic subvariety of $\mathbb{H}^n \times \mathbb{C}^n$.

(ii) If A is an intersection component of $u^{-1}(Z)$ with Zariski closure $\langle A \rangle_{\text{Zar}}$, we define its *Zariski defect* to be

$$\delta_{\text{Zar}}(A) = \dim \langle A \rangle_{\text{Zar}} - \dim A.$$

(iii) An intersection component A of $u^{-1}(Z)$ is called *Zariski optimal* if one cannot find a larger intersection component of $u^{-1}(Z)$ which does not increase the Zariski defect.

(iv) An intersection component A of $u^{-1}(Z)$ is called *geodesic* if A is a component of $u^{-1}(Z) \cap \langle A \rangle_{\text{Zar}}$ and $\langle A \rangle_{\text{Zar}}$ is weakly special.

Proposition 4.2. *Let $Z \subset Y^n \times E^n$ be a subvariety. A Zariski optimal component of $u^{-1}(Z)$ is geodesic.*

Proof. The equivalence of 3.1 and 3.2 is purely formal and the proof is carried out in [13, below 5.12]. ∎

Definition. A *Möbius subvariety* of $\mathbb{H}^n$ is an algebraic subvariety defined by setting some coordinates constant, and relating some other pairs of coordinates by elements of $\mathrm{SL}_2(\mathbb{R})$.

We let F denote a standard fundamental domain for the uniformization of $Y \times E$. The uniformization map restricted to F is definable (in this case by results of Peterzil–Starchenko [24]), and the Möbius subvarieties of $\mathbb{H}^n$ form a definable family.

This means that if we consider the definable family of subvarieties of $\mathbb{H}^n \times \mathbb{C}^n$ comprising all products of Möbius suvarieties of $\mathbb{H}^n$ and linear subvarieties of $\mathbb{C}^n$, and define the set of Zariski optimal ones by the difference of their dimension and the dimension of the intersection with $u^{-1}(V)$, just among these which go through F, we will get elements of $\mathrm{SL}_2(\mathbb{R})$ whose graphs define the slopes (up to Γ and Λ) of all geodesic optimal components. By the *slope* of a Möbius subvariety in $\mathbb{H}^n$ we mean, for each pair of dependent coordinates, the element $g \in \mathrm{SL}_2(\mathbb{R})$ such that the projection to $\mathbb{H}^2$ is the (projection of the) graph of the action by g. And by the *slope* of a weakly special subvariety in $\mathbb{C}^n$, as the uniformizing space for E^n where $E = \Lambda\backslash\mathbb{C}$, we mean the linear subvariety of $\mathbb{C}^n$ which projects to a translate of the weakly special subvariety. This then implies the finiteness of such slopes in $Y^n \times E^n$, and any geodesic optimal component of V^n will have some preimage component going through F.

We need the corresponding finiteness for the particular type of components we consider. Namely, if W is a dependent component, we consider a component U of its preimage in $\mathbb{H}^n \times \mathbb{C}^n$. It is a component of the intersection of $u^{-1}(V^n)$ with suitable preimage $M \times L$ of $\langle W\rangle = S \times B$, and is thus a geodesic component which projects onto M and thus has $\dim U = \dim M$.

We need to observe that, if Zariski optimal, such a component comes from a maximal dependent (weakly) special image, i.e. something of the same form. In fact we need something further along these lines in the proof of 1.1, in order to get from "something positive-dimensional algebraic" to a component of the right form.

Proposition 4.3. *Let U be of the following type: it is a component of $A \times L$ intersecting $u^{-1}(V^n)$, where A is algebraic, and L is linear which projects onto A.*

If U is maximal of this type for the given L then L and A are weakly special and U is Zariski optimal.

Proof. We have $\dim U = \dim A$ and so

$$\delta_{\rm Zar}(U) \le \dim L.$$

Suppose that $U \subset U'$, where U' is Zariski optimal, and hence geodesic optimal, with U' a component of the intersection of $u^{-1}(V^n)$ with weakly special $A' \times L'$, and $A' \times L'$ is the Zariski closure of U'. Then

$$\delta_{\rm Zar}(U') = \dim A' + \dim L' - \dim U'.$$

But $\dim U' \le \dim A'$ and $L \subset L'$. If

$$\delta_{\rm Zar}(U') \le \delta_{\rm Zar}(U)$$

we must have $L = L'$ and $\dim U' = \dim A'$ so that U' is a preimage of a weakly distinguished component. By the maximality of U we have $U = U'$ and then $L = L'$ and $A = A'$ are weakly special. ∎

Now we get the finiteness statement (analogous to [25, Prop. 10.2, and the definable family version 10.3/13.1]). A *strongly special subvariety* in $Y(1)^n$ is defined in [13, 3.2]; it is a fibred product of modular curves. In Y^n it is generally a special subvariety which projects dominantly to each factor.

Proposition 4.4. *For each n there are only finitely many strongly special subvarieties in Y^n which have a V-image which lies in any proper weakly special subvariety of E^n. Moreover, the same holds even if one allows V to vary in a definable family.*

Proof. We prove the stronger version for a definable family of V. We take the definable space of products $M \times L$ of Möbius and linear subvarieties, and take the definable subset of maximal ones in the above sense. As the maximal such varieties are Zariski-optimal, and hence geodesic optimal, the set of slopes of the corresponding Möbius varieties cannot be positive-dimensional, and hence, being definable, it is finite. This set of slopes contains the slopes of all the strongly special subvarieties described in the proposition. ∎

5. Arithmetic estimates

Constants $C, C', \dots, c, c', \dots$ in the following depend on E, Y, V, n and the choice of a fundamental domain F_Y for the uniformization $\mathbb{H} \to Y$. We fix Y to be a *Shimura curve*. In what follows, h, H denote the logarithmic and exponential heights.

For a special point $s \in Y$, we let $\Delta(s)$ denote the discriminant of the corresponding CM abelian variety associated to s by virtue of Y being of abelian type. Recall that in the general case, this is the discriminant of the centre of the endomorphism ring, and in the classical case where Y is the modular curve, it is equal to the discriminant of the corresponding quadratic order.

Proposition 5.1. *Let $s \in Y$ be a special point with discriminant $\Delta(s)$ and let $z \in F_Y$ be a preimage of s. Then*

(1) $h(s) \le c(\epsilon)|\Delta(s)|^{\epsilon}$ *for any* $\epsilon > 0$*;*

(2) $H(z) \le C|\Delta(s)|^{c}$*;*

(3) $[\mathbb{Q}(s) : \mathbb{Q}] \ll \Delta(s)|^{1/2+\epsilon}$ *for any* $\epsilon > 0$*;*

(4) $[\mathbb{Q}(s) : \mathbb{Q}] \gg c(\delta)|\Delta(s)|^{\delta}$ *for some fixed* $\delta > 0$*.*

Proof. For classical singular moduli: (1) Given in [13, Lemma 4.3]. (2) Elementary (with $c = 1$), given in [25]. (3) See [22] for an explicit result. (4) Use the classical (ineffective) Landau–Siegel bound. The same bounds follow for a modular curve Y as a finite cover of $Y(1)$.

For Shimura curves: (2) follows from work of the second author appearing in [26], (1), (3), and (4) follow from [35] combined with the comparison (see e.g. [23]) of Faltings height with height of a moduli point. ∎

We assume E is in Weierstrass form (but an estimate of the same form then follows if it is not) and defined over a number field K_0 of degree $D_0 = [K_0 : \mathbb{Q}]$. Let q denote the Néron–Tate height on E (see e.g. [4] or [16]).

We have the following Theorem E of Masser [16]. Set

$$\eta = \eta(E, K) = \inf q(x),$$

with the infimum taken over non-torsion $x \in E(K)$, and let

$$\omega = \omega(E, K)$$

be the cardinality of the torsion subgroup of $E(K)$.

Theorem 5.2. *Let $x_1, \dots, x_n \in E(K)$ with Néron–Tate heights bounded by $q \geq \eta$. There is a basis for the relations*

$$m_1 x_1 + \cdots + m_n x_n = 0_E, \quad m_i \in \mathbb{Z},$$

with all m_i having

$$|m_i| \leq n^{n-1} \omega (q/\eta)^{(n-1)/2}. \qquad \blacksquare$$

To accommodate CM, we work, like Barroero [2], in E^{2n} with $x_i, \rho x_i$, where E has CM by the order $\mathbb{Z} + \mathbb{Z}\rho$. We have $q(\rho x) = N(\rho) q(x) = |\rho|^2 q(x)$ (see e.g. [34, Lemma 1]). We write $\|a + b\rho\| = \max\{|a|, |b|\}$ for $a + b\rho \in \mathrm{End}(E)$. Then under the previous hypotheses a set of generators for the relation group can be found with

$$\|m_i\| \leq (\max\{1, |\rho|\})^{2n-1} (2n)^{2n-1} \omega (q/\eta)^{(2n-1)/2},$$

where q is an upper bound for $q(x_i)$, $i = 1, \dots, n$.

Following [16] we have the following estimates for η, ω, where we set $D = [K : \mathbb{Q}]$ and $L = \log(D + 2)$:

$$\eta \geq C^{-1} D^{-3} L^{-2}$$

by results of, respectively, Laurent (CM) and Masser (non-CM) cited in [16], and

$$\omega \leq CDL$$

(see discussion in [16]).

Combining the above estimates yields the following result, where $\|m\|$ is as above in the CM case, but in the non-CM case we set $\|m\| = |m|$.

For a tuple $s = (s_1, \dots, s_n) \in Y^n$ of special points with discriminants $\Delta(s_i)$ we define the *complexity* of s by $\mathbf{\Delta}(s) = \max |\Delta(s_i)|$.

Proposition 5.3. *There are constants C, C', c, depending on E, Y, V, n, with the following property. Let $(s_1, x_1) \dots, (s_n, x_n) \in Y \times E$ be V-graphs of special points with discriminants $\Delta(s_i)$ and set $\mathbf{\Delta} = \mathbf{\Delta}(s) = \mathbf{\Delta}(s_1, \dots, s_n)$. Then, for $\mathbf{\Delta} \geq C'$, there is a generating set for the linear relations satisfied by the x_i in E with*

$$\|m_i\| \leq C \mathbf{\Delta}^c.$$

Proof. The difference $|q - h|$ is bounded on $E(\overline{K_0})$ by some constant c^* (see e.g. [4]). On the other hand, if x is a V-image of s then $H(x) \leq CH(s)^c$ and

$$[K_0(x) : K_0] \leq C[\mathbb{Q}(s) : \mathbb{Q}].$$

Thus, $D \leq C\mathbf{\Delta}^c$ by 5.1(3).

If the maximum h of the $h(x_i)$ is sufficiently large then we will have $h - c^* \geq \eta$ and $2h \geq q$. Then $h \leq C\mathbf{\Delta}^c$ by 5.1(1), and now all the constituents of the bound in 5.2 are bounded in terms of $\mathbf{\Delta}$. $\blacksquare$

Propositions 5.3 and 5.1(2) will be used in the next section to bound the height of a rational/quadratic point on a suitable definable set, while 5.1(4) will be used to show that there are “many” such points.

6. Proof of theorems over $\overline{\mathbb{Q}}$

Proof of Theorem 1.1 *when E, V are defined over* $\overline{\mathbb{Q}}$. Let K_0 be a number field over which E, Y, V and all elements of $\mathrm{End}(E)$ are defined.

We consider an exemplary component $W \subset V^n$, a V-image of some special subvariety $S \subset Y^n$, with $\langle \pi_{E^n}(W) \rangle = B$. Then any Galois conjugate W' of W over K_0 is also an exemplary component (of the conjugate S' of S, with $\langle \pi_{E^n}(W') \rangle = B'$ and B' the corresponding conjugate of B), and vice versa.

By Lemma 2.1 we can write S as a product $S = S_1 \times \{S_2\}$ of some strongly special $S_1 \subset Y^{A_1}$ on some subset $A_1 \subset \{1, \dots, n\}$ of coordinates, and a special point $S_2 \in Y^{A_2}$ where $A_2 \subset \{1, \dots, n\}$ is the complement subset to A_1.

By Proposition 3.4 there are only finitely many such S_1 to consider, and so we may assume they are all defined over K_0.

We can write $W = W_1 \times W_2$ and write ξ_j, η_k for the coordinates in E^{A_1}, E^{A_2} respectively. We will show that if $\eta \in E^{A_2}$ is a V^{A_2}-image of a special point S_2 of sufficiently large complexity (depending on S_1) then W is not exemplary, and this will establish the requisite finiteness.

It may be that the projection of W_1 to E^{A_1} is contained in some proper weakly special subvariety, which means that there are some equations of the form

$$\sum_{i \in A_1} m_i \xi_i = p, \quad m_i \in \mathrm{End}(E),\ p \in E,$$

holding on this projection. We let $p_1, \dots, p_k$ be the points corresponding to a generating set of such relations. Note that the linear span of the p_i is $\mathrm{Gal}(\overline{\mathbb{Q}}/K_0)$-invariant, so we can make all the p_i defined over K_0.

If we take a generating set of all the equations over $\mathrm{End}(E)$ satisfied by the points in $\pi_{E^n}(W)$ then this defines an algebraic subgroup B_0 of which B is a connected component. Any such equation of the form

$$\sum_{i \in A_1} m_i \xi_i + \sum_{j \in A_2} n_j \eta_j = 0, \quad m_i, n_j \in \mathrm{End}(E),$$

entails that $\sum m_i \xi_i$ is constant on W_1 and is equivalent to some equation involving the p_i, η_j, and vice versa. We then consider the system of equations

$$\sum_{i=1}^{k} m'_i p_i + \sum_{j \in A_2} n_j \eta_j = 0, \quad m'_i, n_j \in \mathrm{End}(E),$$

corresponding (and equivalent) to the system defining B_0, where η is a V^{A_2}-image of S_2. Let d_0 be the dimension of the subvariety this cuts out in E^{A_2}.

By Proposition 5.3 there is a set of generators of all such relations with

$$\|m_i\|, \|n_j\| \le C\, \mathbf{\Delta}(S_2)^c.$$

Fix a preimage $\nu = (\nu_1, \ldots, \nu_k) \in F_E^k$ of $(p_1, \ldots, p_k)$. Let us first suppose that E has NCM (“not CM”), and $d = \dim B$. Let G be the Grassmannian of $(d_0 + k)$-dimensional affine linear $\mathbb{C}$-subspaces of $\mathbb{C}^{k+n_2}$ where $n_2 = |A_2|$.

Take the definable set

$$X = \{(z, w, g) \in F_Y^{A_2} \times F_E^{A_2} \times G : u(z, w) \in V^{A_2},\ (\nu, w) \in g\},$$

where F_E is a standard fundamental domain for the uniformization $\mathbb{C} \to E$, and, projecting, the definable set

$$Z = \{(z, g) \in F_Y^{A_2} \times G : \exists w \in F_E^{A_2} : (z, w, g) \in X\}.$$

A special point $S_2 \in Y^{A_2}$ of “large” complexity $\mathbf{\Delta}(s)$ leads to “many” points in Z which are quadratic in the F_Y coordinates and rational (even integral) in the g coordinates. More specifically, for sufficiently large $\mathbf{\Delta}(s)$ we get (by 5.1(4), 5.1(2), and 5.3)

$$\gg \mathbf{\Delta}(S_2)^c \text{ such points of height at most } \ll \mathbf{\Delta}(S_2)^{c'}.$$

Hence, by the Counting Theorem (see e.g. [28]), there is a connected, semialgebraic set R in Z belonging to a fixed definable family, in which the z coordinates cannot be constant (since the positive-dimensional semialgebraic sets need to account for “many” different conjugates of s). Since all of the Galois conjugates of a point have the same slopes m_i, n_j we can moreover assume that R has a fixed slope.

Lemma 6.1. *The projection of R to G is a point.*

Proof. Let β be the covering space of B_0 and $\beta' = \mathbb{C}^{A_2}/\beta$. Consider the image $R' \subset F_Y^{A_2} \times F_{\beta'}$ of the preimage of R in X. Again by the counting theorem, R' contains a semialgebraic set R'' belonging to a fixed definable family, with “many" rational points coming from a single Galois orbit. Now note that R'' maps into the image V' of V^n inside the product $Y^{A_2} \times E^{A_2}/B_0$. Thus by Ax–Lindemann, the image of R'' lies in a weakly special subvariety contained in V'. However, the projection of V' to Y^{A_2} is finite-to-one, and therefore the weakly special subvariety containing the image of R'' must have no abelian part, and therefore its projection to E^{A_2}/B_0 is a point, as desired. ∎

By Lemma 6.1 we may write $R = A \times g_0$ with $g_0 \in G$ and $A \subset \mathbb{H}^{A_2}$ semialgebraic. Let L be the linear subspace of $\mathbb{C}^{k+n_2}$ corresponding to g_0. Note that L projects to some Galois conjugate of B inside E^{A_2}. Let $L_\nu \subset \mathbb{C}^k$ be the fibre of L over ν. Now, by definition of A, we know that $A \times L_\nu \cap u^{-1}(V^{A_2})$ has a component U which maps onto A. Note that the Zariski defect of U is at most d_0.

By Proposition 4.3, there exists a weakly special subvariety A^* containing A and a component U^* of $A^* \times L_\nu \cap u^{-1}(V^{A_2})$ containing U which maps onto A^* with defect at most d_0. Since A^* contains special points, it must in fact be special. Let S^* be the image of A^* in Y^{A_2}. It contains at least one (in fact “many”) Galois conjugates of S_2. By definition, a suitable V-image of $S_1 \times S^*$ is contained in a coset of B_0. We may now take a Galois conjugate of S^* which contains S_2, thus giving a larger exemplary component projecting to the same torsion coset, which is a contradiction.

Now suppose that E has CM by the order $\mathbb{Z}+\mathbb{Z}\rho$. We now let G parameterize (n_2+2k+d_0)-dimensional complex affine subspaces in $\mathbb{C}^{2k+2n_2}$ and consider the definable set

$$X=\{(z,w,g)\in F_Y^{A_2}\times F_E^{A_2}\times G: u(z,w)\in V^{A_2}, (\nu,\rho\nu,w,\rho w)\in g\}$$

and, projecting, the definable set

$$Z=\{(z,g)\in F_Y^{A_2}\times G: \exists w\in F_E^{A_2}: (z,w,g)\in X\}.$$

The rest of the proof is as in the NCM case. ■

Proof of Theorem 1.5 *when* E, V, U *are defined over* $\overline{\mathbb{Q}}$. This is very much the same as the argument above but using different arithmetic estimates, drawn from [12], and a different definable set on which to count points.

We again consider an exemplary U-component of the form $W_1\times W_2$, a V-image of some $S_1\times\{S_2\}$ as above with $S_2\in Y^{A_2}$ a U-special point. There are again only finitely many such decompositions to consider, by 4.4.

Let us consider U-special points $S_2=(s_i)\in Y^{A_2}$ of a particular form, namely points in which s_i is in the Hecke orbit of a fixed $u_i\in U$ for $i\in A_2$, and all the u_i are non-special. Then there is a unique cyclic isogeny between the elliptic curves corresponding to u_i and s_i whose degree we denote N_i. For such a point S_2 we define its *U-complexity* by

$$\mathbf{\Delta}(S_2)=\max\{N_1,\ldots,N_n\}.$$

We observe that the height of S_2 is controlled by $\mathbf{\Delta}(S_2)$; using the results of Faltings relating Faltings heights of isogenous elliptic curves and Silverman's comparison of Faltings height and height of the j-invariant (see the discussion on heights under isogenies in [12, proof of Lemma 4.2, p. 15]) we have

$$h(S_2)\le C\max\{1,\log N_i\}$$

(constants now depend on Y, E, V, U and n). If $(S_2,\eta)\in V^{A_2}$, the above leads (via Masser's Theorem E) to bounds of the form

$$\|m\|\le C\,\mathbf{\Delta}(S_2)^c$$

on the size of entries in a set of generators for the relation group of (p,η).

On the other hand, the degrees $[\mathbb{Q}(s_i):\mathbb{Q}]$ are controlled by $\mathbf{\Delta}(S_2)$ via isogeny estimates (see the discussion on degrees in [12, §6, above proof of 1.3]) which imply $[\mathbb{Q}(s_i):\mathbb{Q}]\ge C'N_i^{1/6}$ and hence

$$[\mathbb{Q}(S_2):\mathbb{Q}]\ge C'\mathbf{\Delta}(S_2)^{c'}.$$

Finally, if $\nu_i\in F_Y$ is a preimage of u_i and $z_i\in F_Y$ is a preimage of s_i then $z_i=g\nu_i$ for some $g_i\in \mathrm{GL}_2^+(\mathbb{Q})$ with

$$H(g_i)\le cN_i^{10}$$

(see [12, Lemma 5.2]).

We now count points though on a different definable set, because U-special points are not algebraic and the counting must be done for $\mathrm{GL}_2^+(\mathbb{Q})$ points in a definable subset of $\mathrm{GL}_2^+(\mathbb{R})$.

We fix a preimage $\mu \in F_Y^{A_2}$ of $(u_1, \ldots, u_{n_2})$ and consider the definable set

$$X = \{(h, \nu, w, g) \in \mathrm{GL}_2^+(\mathbb{R})^{A_2} \times F_E^k \times F_E^{A_2} \times G : \\ h\mu \in F_Y^{A_2},\ u(h\mu, w) \in V^{A_2},\ (\nu, w) \in g\}$$

and its projection

$$Z = \{(h, g) \in \mathrm{GL}_2^+(\mathbb{R})^{A_2} \times G : \exists w \in F_E^{A_2} \colon (h\mu, \nu, w, g) \in X\}.$$

A U-special point S_2 of the form being considered of "large" complexity leads to "many" rational points on Z. If $\mathbf{\Delta}(S_2)$ is sufficiently large then by counting we get a real algebraic curve in Z which (since these come from "many" distinct points in $F_Y^{A_2}$ and by complexification) gives rise to a complex algebraic curve $A \subset \mathbb{H}^{A_2}$ and an intersection component of $A \times L_g$ of Zariski defect d as previously. This leads to a contradiction as in the argument above, so that $\mathbf{\Delta}(S_2)$ is bounded for an exemplary U-component, giving finiteness for S_2 of this type.

The general case will follow by combining the treatment of special and non-special points using a suitable definable set (i.e. using F_Y for special coordinates and $\mathrm{GL}_2^+(\mathbb{R})$ for coordinates in the Hecke orbit of a non-special $u \in U$) and a combinatorial argument. ∎

7. Going from $\overline{\mathbb{Q}}$ to $\mathbb{C}$

We fix Y to be a Shimura curve, and a positive integer t. Consider the following statement:

Conjecture 7.1. *Let E be an elliptic curve. Given $V \subset Y \times E$, $U \subset Y$ finite of size t, and $n \geq 1$, there are only finitely many exemplary U-components in V^n.*

We shall prove in this section that Conjecture 7.1 over $\overline{\mathbb{Q}}$ implies it in general. This, combined with the work in the previous section, proves Theorems 1.1 and 1.5.

7.1. Setup

Let F be a finitely generated subfield of $\mathbb{C}$ such that $V \subset Y(1) \times E$ and U are all defined over F. Then F can be thought of as the function field of an irreducible algebraic variety S over some number field $K \subset F$. Replacing S with a dense open subset, we assume that E extends to an elliptic scheme $\mathcal{E}$ over S and V extends to a flat family $\mathcal{V}$ over S.

We pick a generic regular point $s_0 \in S(\mathbb{C})$ such that $K(s_0)$ is isomorphic to F, and pick an open ball $B \subset S(\mathbb{C})$ around s_0, so that in B we can trivialize the homology of $\mathcal{E}$ over S.

7.2. *Ordering points in S*

We will need to order points in S, so we proceed as follows. Let $f : S \to \mathbb{P}^{\dim S}$ be a quasi-finite map. Then we define the *f-degree* of a point s in $S(\overline{\mathbb{Q}})$ to be the degree of its image under f, and the *f-height* $h_f(s)$ to be the (logarithmic) height of its image under f. By Northcott's theorem, there are finitely many points of bounded f-degree and f-height. We only consider heights for the subset S_f of S whose image lands in $\mathbb{P}^{\dim S}(K)$.

7.3. *The proof that Conjecture 7.1 over $\overline{\mathbb{Q}}$ implies Conjecture 7.1 over $\mathbb{C}$*

By Proposition 4.4 there are only finitely many strongly special subvarieties whose $\mathcal{V}_u$-image lies inside any proper weakly special subvariety of $\mathcal{E}_u$ for any $u \in B$. Thus, there are only finitely many families of special subvarieties we have to consider. By rearranging coordinates, we may assume they are all of the form $T \times p \times q$ where $T \subset Y^m$ is a fixed strongly special subvariety, and $p \in Y(1)^k$ is a CM point, and q has coordinates isogenous to points in U.

Now, for the sake of contradiction let p_i, q_{i,s_0} be an infinite sequence of such points such that $T \times p_i \times q_{i,s_0}$ are projections of U-exemplary components for $\mathcal{V}_{s_0}$. Let A_i be the smallest torsion coset containing the $\mathcal{V}_{s_0}$-image of $T \times p_i \times q_{i,s_0}$. Then for each point $s \in S(\overline{\mathbb{Q}}) \cap B$, the image of $T \times p_i \times q_{i,s}$ is still contained in $A_{i,s}$. But we have assumed the statement for $\overline{\mathbb{Q}}$-points, and thus for each s there are finitely many U-special subvarieties containing all the $T \times p_i \times q_{i,s}$ whose $\mathcal{V}_s$-image is contained in a proper torsion coset.

Let $T_1(s_0), \dots, T_m(s_0)$ be the smallest collection of U_{s_0}-special subvarieties containing all the $T \times p_i \times q_{i,s_0}$.

Lemma 7.1. *For large enough d, for a density 1 set of points s in S_f ordered by f-height, $T_1(s), \dots, T_m(s)$ is the smallest collection of U_s-special subvarieties containing all the $T \times p_i \times q_{i,s}$.*

Proof. First, note that the degrees of CM points tend to infinity. Thus, the set of points $s \in S_f$ such that U_s is CM is contained in a proper subvariety, and so has density 0. Next, since U-special subvarieties are defined simply by imposing isogeny relations, it is sufficient to prove that for a density 1 set of points s the points u_s, v_s are not isogenous, for u, v distinct points in U.

Now, for $s \in S_f$, it follows that $h(u_s), h(v_s) \ll h_f(s)$, and thus by the Masser–Wüstholz isogeny bound [18, Main Theorem] it follows that if u_s, v_s are isogenous then there is an isogeny between them of degree $O(h_f(s)^\kappa)$ for some fixed $\kappa > 0$. Now, the degree of the N-isogeny releation in Y^2 is $O(N^\eta)$ for some fixed $\eta > 0$, and therefore the set of all $s \in S_f$ with $h_f(s) < X$ such that u_s, v_s are isogenous is contained in $O(X^\kappa)$ divisors of f-degree at most $O(X^\eta)$. Now, the size of $\{s \in S_f : h(s) < X\}$ is asymptotic to $e^{[K:\mathbb{Q}]X(\dim S+1)}$, whereas the number of points in any divisor of degree d of height at most X is $O(de^{[K:\mathbb{Q}]X \dim S})$. The result follows. ∎

Thus we are done once we prove the following.

Lemma 7.2. *Let $\mathcal{E}$ be an elliptic scheme over S, and let $\mathcal{W} \subset \mathcal{E}^n$ be an irreducible algebraic subvariety. If $\mathcal{W}_s$ is contained in a proper abelian subvariety for a density* 1 *set of $s \in S_f$, then $\mathcal{W}$ is contained in a proper abelian subscheme.*

Proof. Replacing $\mathcal{W}$ by its own n-fold self-sum we may assume that $\mathcal{W}$ is a coset of an abelian subscheme. Quotienting out by the corresponding abelian subscheme, we may further assume that $\mathcal{W}$ is finite over S, and base changing S by a finite map we may assume that $\mathcal{W}$ is a section over S. By the Main Theorem of [17], it follows that for a density 1 set of points s the n points of $\mathcal{E}_s$ represented by $\mathcal{W}_s$ are linearly independent. This completes the proof when $\mathcal{E}$ does not have generic CM. If $\mathcal{E}$ has generic CM, we let ρ be an extra endomorphism. Then we conclude by applying the same theorem of [17] to the $2n$ points of $\mathcal{E}_s$ given by the coordinates of $\mathcal{W}_s$ as well as those of $\mathfrak{z}\mathcal{W}_s$. ∎

Acknowledgements. The authors are grateful to the referees for pointing out several inaccuracies and obscurities.

Funding. JP thanks EPSRC for support under grant reference EP\N008359\1. JT thanks NSERC and the Ontario Early Researcher Award for support.

References

[1] Baldi, G.: On a conjecture of Buium and Poonen. Ann. Inst. Fourier (Grenoble) **70**, 457–477 (2020) Zbl 1450.11059 MR 4105944

[2] Barroero, F.: CM relations in fibered powers of elliptic families. J. Inst. Math. Jussieu **18**, 941–956 (2019) Zbl 1455.11080 MR 3995717

[3] Bombieri, E., Masser, D., Zannier, U.: Anomalous subvarieties—structure theorems and applications. Int. Math. Res. Notices **2007**, art. rnm057, 33 pp. Zbl 1145.11049 MR 2359537

[4] Bruin, P.: Bornes optimales pour la différence entre la hauteur de Weil et la hauteur de Néron–Tate sur les courbes elliptiques sur $\overline{\mathbb{Q}}$. Acta Arith. **160**, 385–397 (2013) Zbl 1287.11083 MR 3119786

[5] Buium, A., Poonen, B.: Independence of points on elliptic curves arising from special points on modular and Shimura curves. I. Global results. Duke Math. J. **147**, 181–191 (2009) Zbl 1177.11050 MR 2494460

[6] Buium, A., Poonen, B.: Independence of points on elliptic curves arising from special points on modular and Shimura curves. II. Local results. Compos. Math. **145**, 566–602 (2009) Zbl 1232.11073 MR 2507742

[7] Daw, C., Ren, J.: Applications of the hyperbolic Ax–Schanuel conjecture. Compos. Math. **154**, 1843–1888 (2018) Zbl 1427.11057 MR 3867286

[8] Deligne, P.: Travaux de Shimura. In: Séminaire Bourbaki, 23ème année (1970/71), Lecture Notes in Math. 244, exp. 389, Springer, Berlin, 123–165 (1971) Zbl 0225.14007 MR 0498581

[9] Dill, G. A.: Unlikely intersections with isogeny orbits in a product of elliptic schemes. Math. Ann. **377**, 1509–1545 (2020) Zbl 07229164 MR 4126900

[10] Dill, G. A.: Unlikely intersections between isogeny orbits and curves. J. Eur. Math. Soc. **23**, 2405–2438 (2021) Zbl 07367662 MR 4269417

[11] Gao, Z.: Mixed Ax–Schanuel for the universal abelian varieties and some applications. Compos. Math. **156**, 2263–2297 (2020) Zbl 07305885 MR 4190045

[12] Habegger, P., Pila, J.: Some unlikely intersections beyond André–Oort. Compos. Math. **148**, 1–27 (2012) Zbl 1288.11062 MR 2881307

[13] Habegger, P., Pila, J.: O-minimality and certain atypical intersections. Ann. Sci. École Norm. Sup. (4) **49**, 813–858 (2016) Zbl 1364.11110 MR 3552014

[14] Khare, C., Rajan, C. S.: On Heegner points of large conductors. Math. Res. Lett. **8**, 275–278 (2001) Zbl 1067.11033 MR 1839477

[15] Kühne, L.: Intersection of class fields. Acta Arith. **198**, 109–127 (2021) Zbl 07349951 MR 4228297

[16] Masser, D. W.: Linear relations on algebraic groups. In: New Advances in Transcendence Theory (Durham, 1986), Cambridge Univ. Press, Cambridge, 248–262 (1988) Zbl 0656.10031 MR 972004

[17] Masser, D. W.: Specializations of finitely generated subgroups of abelian varieties. Trans. Amer. Math. Soc. **311**, 413–424 (1989) Zbl 0673.14016 MR 974783

[18] Masser, D. W., Wüstholz, G.: Some effective estimates for elliptic curves. In: Arithmetic of Complex Manifolds (Erlangen, 1988), Lecture Notes in Math. 1399, Springer, Berlin, 103–109 (1989) Zbl 0701.14034 MR 1034259

[19] Mok, N., Pila, J., Tsimerman, J.: Ax–Schanuel for Shimura varieties. Ann. of Math. (2) **189**, 945–978 (2019) Zbl 07107179 MR 3961087

[20] Moonen, B.: Linearity properties of Shimura varieties. I. J. Algebraic Geom. **7**, 539–567 (1998) Zbl 0956.14016 MR 1618140

[21] Nekovář, J., Schappacher, N.: On the asymptotic behaviour of Heegner points. Turkish J. Math. **23**, 549–556 (1999) Zbl 0977.11025 MR 1780940

[22] Paulin, R.: An explicit André–Oort type result for $\mathbb{P}^1(\mathbb{C}) \times \mathbb{G}_m(\mathbb{C})$ based on logarithmic forms. Publ. Math. Debrecen **88**, 21–33 (2016) Zbl 1374.11070 MR 3452162

[23] Pazuki, F.: Theta height and Faltings height. Bull. Soc. Math. France **140**, 19–49 (2012) Zbl 1245.14029 MR 2903770

[24] Peterzil, Y., Starchenko, S.: Definability of restricted theta functions and families of abelian varieties. Duke Math. J. **162**, 731–765 (2013) Zbl 1284.03215 MR 3039679

[25] Pila, J.: O-minimality and the André–Oort conjecture for $\mathbb{C}^n$. Ann. of Math. (2) **173**, 1779–1840 (2011) Zbl 1243.14022 MR 2800724

[26] Pila, J., Tsimerman, J.: The André–Oort conjecture for the moduli space of abelian surfaces. Compos. Math. **149**, 204–216 (2013) Zbl 1304.11055 MR 3020307

[27] Pila, J., Tsimerman, J.: Multiplicative relations among singular moduli. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (5) **17**, 1357–1382 (2017) Zbl 1408.14092 MR 3752529

[28] Pila, J., Wilkie, A. J.: The rational points of a definable set. Duke Math. J. **133**, 591–616 (2006) Zbl 1217.11066 MR 2228464

[29] Pink, R.: A combination of the conjectures of Mordell–Lang and André–Oort. In: Geometric Methods in Algebra and Number Theory, Progr. Math. 235, Birkhäuser Boston, Boston, MA, 251–282 (2005) Zbl 1200.11041 MR 2166087

[30] Pink, R.: A common generalization of the conjectures of André–Oort, Manin–Mumford, and Mordell–Lang. Manuscript dated 17 April 2005, http://www.math.ethz.ch/$\sim$pink/

[31] Poizat, B.: L'égalité au cube. J. Symbolic Logic **66**, 1647–1676 (2001) Zbl 1005.03038 MR 1878287

[32] Rosen, M., Silverman, J. H.: On the independence of Heegner points associated to distinct quadratic imaginary fields. J. Number Theory **127**, 10–36 (2007) Zbl 1151.11024 MR 2351661

[33] Şahinoğlu, H.: On the independence of Heegner points on CM curves associated to distinct quadratic imaginary fields. Proc. Amer. Math. Soc. **141**, 813–826 (2012) Zbl 1294.11090 MR 3003675

[34] Siksek, S.: The height pairing on elliptic curves with complex multiplication. Arab J. Math. Sci. **5**, 43–48 (1999) Zbl 0991.11027 MR 1698566

[35] Tsimerman, J.: The André–Oort conjecture for $\mathcal{A}_g$. Ann. of Math. (2) **187**, 379–390 (2018) Zbl 1415.11086 MR 3744855

[36] Zilber, B.: Exponential sums equations and the Schanuel conjecture. J. London Math. Soc. (2) **65**, 27–44 (2002) Zbl 1030.11073 MR 1875133