COMPOSITIO MATHEMATICA

Some unlikely intersections beyond André–Oort

P. Habegger and J. Pila

Compositio Math. **148** (2012), 1–27.

doi:10.1112/S0010437X11005604

FOUNDATION
COMPOSITIO
MATHEMATICA

LONDON
MATHEMATICAL
SOCIETY

Compositio Math. 148 (2012) 1–27
doi:10.1112/S0010437X11005604

Some unlikely intersections beyond André–Oort

P. Habegger and J. Pila

Abstract

According to the André–Oort conjecture, an algebraic curve in $Y(1)^n$ that is not equal to a special subvariety contains only finitely many points which correspond to an n-tuple of elliptic curves with complex multiplication. Pink's conjecture generalizes the André–Oort conjecture to the extent that if the curve is not contained in a special subvariety of positive codimension, then it is expected to meet the union of all special subvarieties of codimension two in only finitely many points. We prove this for a large class of curves in $Y(1)^n$. When restricting to special subvarieties of codimension two that are not strongly special we obtain finiteness for all curves defined over $\overline{\mathbb{Q}}$. Finally, we formulate and prove a variant of the Mordell–Lang conjecture for subvarieties of $Y(1)^n$.

1. Introduction

We verify a particular case of Pink's generalization of the André–Oort conjecture [Pin05] regarding unlikely intersections of an algebraic subvariety of a Shimura variety with varying special subvarieties. As explained below, our results fit into the framework of conjectures on 'unlikely intersections' formulated first by Zilber, in the semi-abelian setting, and, most generally, by Pink for mixed Shimura varieties.

Let n be a positive integer. Our ambient Shimura variety will be a power $Y(1)^n$ of the modular curve $Y(1)$, the latter being the affine line. The complex points of $Y(1)$ can be identified with isomorphism classes of elliptic curves defined over $\mathbb{C}$.

The *special subvarieties of* $Y(1)^n$, or just *special subvarieties* for short, play an important role in our results. We refer to § 2.1 for a precise definition and now only give an informal description. We consider an n-tuple of elliptic curves whose j-invariants constitute the coordinates of a point on a special subvariety of $Y(1)^n$. If the special subvariety has dimension zero, then the elliptic curves all have complex multiplication and the corresponding tuple of j-invariants is called a *special point*. For a general special subvariety, certain elliptic curves may have complex multiplication and certain pairs may be isogenous.

Let m be a non-negative integer. We set

$$\mathcal{S}^{[m]} = \bigcup_{\substack{S \subset Y(1)^n \\ \operatorname{codim} S \geqslant m}} S(\mathbb{C})$$

where the union runs over all special subvarieties S of $Y(1)^n$ of codimension at least m.

The André–Oort conjecture is a statement on Shimura varieties which governs the intersection of a subvariety with the set of special points. Roughly speaking, it states that the special subvarieties are exactly those subvarieties that contain a Zariski dense set of special points.

Received 14 September 2010, accepted in final form 17 March 2011, published online 30 August 2011.
2010 Mathematics Subject Classification 11G18 (primary), 03C64, 11G15, 11G50, 14G35, 14K22 (secondary).
Keywords: André–Oort conjecture, Zilber–Pink conjecture, unlikely intersection.

In our notation, the set of special points in $Y(1)^n$ is $\mathcal{S}^{[n]}$. For curves in $Y(1)^2$ the André–Oort conjecture follows unconditionally from the work of André [And98]. The analogue statement for $Y(1)^n$ follows by projecting to all pairs of two distinct coordinates and from the characterization of special subvarieties of $Y(1)^n$ due to Edixhoven [Edi05]. Edixhoven [Edi98] proved the same result under the Generalized Riemann Hypothesis (GRH), and later [Edi05] established the generalization to arbitrary subvarieties of $Y(1)^n$, also under the GRH. A different proof of the latter result, still under GRH, is due to Ullmo and Yafaev [UY09]. The second author has recently found an unconditional proof of this result [Pil11]. His approach relies on a counting result due to Wilkie and himself [PW06], and adopts a basic strategy proposed by Zannier in the context of the Manin–Mumford conjecture (see [PZ08]); this strategy also plays a central role in the current article.

We will not state the general version of Pink's conjecture 1.3 [Pin05] here but rather formulate it in the particular case of curves in $Y(1)^n$.

Conjecture. Let $C \subset Y(1)^n$ be an irreducible curve defined over $\mathbb{C}$. If C is not contained in a special subvariety of positive codimension, then $C(\mathbb{C}) \cap \mathcal{S}^{[2]}$ is finite.

Further down we will briefly recount some related conjectures and results on unlikely intersections in the different setting of semi-abelian varieties.

Our first result goes in the direction of this conjecture. We are able to handle curves defined over $\overline{\mathbb{Q}}$, an algebraic closure of $\mathbb{Q}$, that satisfy an additional restriction which we now describe.

Let $C \subset Y(1)^n$ be an irreducible curve defined over $\mathbb{C}$ and let $X_1, \dots, X_n$ denote the coordinate functions on $Y(1)^n$. We call C *asymmetric* if any non-zero integer appears at most once in the sequence $\deg(X_1|_C), \dots, \deg(X_n|_C)$ of degrees, up to one exception, which may appear twice. Note that we allow coordinate functions to be constant on C.

Theorem 1. *Let $C \subset Y(1)^n$ be an irreducible curve defined over $\overline{\mathbb{Q}}$. If C is not contained in a special subvariety of positive codimension and if C is asymmetric, then $C(\mathbb{C}) \cap \mathcal{S}^{[2]}$ is finite.*

We remark that any irreducible curve in $Y(1)^2$ is asymmetric. Hence the André–Oort conjecture for curves in $Y(1)^2$ is a formal consequence of our first theorem.

We believe that Theorem 1 continues to hold if C is merely defined over $\mathbb{C}$. A possible approach to this question could involve ideas used by Bombieri *et al.* [BMZ08] who treated curves in the algebraic torus.

The proof of Theorem 1 involves three parts which we discuss briefly. Let C be as in the theorem. In this sketch we suppose that $C(\overline{\mathbb{Q}}) \cap \mathcal{S}^{[2]}$ is infinite and derive a contradiction.

Essential arithmetic information comes from the action of the absolute Galois group of a field of definition of C on $C(\overline{\mathbb{Q}}) \cap \mathcal{S}^{[2]}$. Say x is in this intersection. The first part consists in finding a lower bound for the cardinality of the Galois orbit of x in terms of the 'complexity' of a codimension-two special subvariety containing this point. This is done in § 4. Since no coordinate of x need be the j-invariant of an elliptic curve with complex multiplication, we cannot rely on class field theory to find many conjugates of x. An important tool is a weak height upper bound for x in the spirit of [Hab10]. A new aspect in this paper is that this bound is combined with isogeny estimates for elliptic curves. These appeared prominently in the work of Masser and Wüstholz [MW90]. For our purposes we will need a slightly more explicit version which follows from work of David [Dav95] or of Pellarin [Pel01]. It is at this stage where we need the supplementary hypothesis that C is asymmetric. Substantially stronger upper bounds in these

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

isogeny estimates would imply Theorem 1 without the condition on asymmetry. However, results of the required quality seem out of reach at the moment.

The second part of the proof is contained in § 5. We use an extension [Pil11] of the counting result of Pila–Wilkie which we state precisely in § 2.3. It is formulated in terms of definable sets in an o-minimal structure, a notion we will also recall, and deals with the distribution of algebraic points of fixed degree and bounded height on such sets. We proceed by constructing a suitable definable set using C and the modular j-function. A point such as x above will give rise to an algebraic point z of degree at most two over $\mathbb{Q}$ on the definable set. For example, if x is a special point then z has imaginary quadratic coordinates and the definable set in question is roughly the pre-image of $C(\mathbb{C})$ under the j-function. In the general case, z also involves matrices defining the special subvariety containing x and the definable set resembles an incidence set. At the heart of the strategy lies the observation that any point in the Galois orbit of x leads to a rational point on said definable set. The height of the point z can be related to the 'complexity' of the special subvariety containing x. Essentially, the Galois orbit in the first part leads to an excessive number of such points on the definable set if the 'complexity' is large enough. The counting result implies that this can only happen if the definable set contains a rather large semi-algebraic set. This is at odds with the transcendental nature of the definable set and therefore implies a strong restriction on C. We formulate this here precisely in terms of the nth Cartesian power of the modular j-function $j : \mathbb{H}^n \to Y(1)^n$; here $\mathbb{H}$ is the upper half-plane. It turns out that a local inverse of j restricted to $C(\mathbb{C})$ takes values in an algebraic hypersurface of $\mathbb{C}^n \supset \mathbb{H}^n$.

The third and final part of the proof consists in handling this situation. This is done in § 3. Indeed, here we will conclude that C is either contained in a special subvariety of positive codimension or some $X_i|_C$ is constant. The first case contradicts our hypothesis and the second can be disposed of without much difficulty if the constant coordinate does not attain a special value. The principle statement in this step is given by Proposition 3.1, which relies on a Hodge-theoretic result of André [And92]. This proposition is related to Ax's theorem [Ax71]. Ax proved a variant, involving formal power series, of Schanuel's conjecture on transcendental properties of the exponential function. Our proposition recovers a weak analogue of Ax's theorem for the modular j-function.

The hypothesis on asymmetry is only needed when x lies in a *strongly special subvariety* of codimension two; a strongly special subvariety is a special subvariety on which no coordinate function X_i is constant. Below we will obtain finiteness as in Theorem 1 for curves which need not be asymmetric after omitting the strongly special subvarieties from $\mathcal{S}^{[2]}$. So let us set

$$\mathcal{S}^{\mathrm{nss},[m]} = \bigcup_{\substack{S \subset Y(1)^n \\ \operatorname{codim} S \geqslant m}} S(\mathbb{C}) \subset \mathcal{S}^{[m]}$$

where the union runs over all special subvarieties S of $Y(1)^n$ of codimension at least m that are *not* strongly special.

Theorem 2. *Let $C \subset Y(1)^n$ be an irreducible curve defined over $\overline{\mathbb{Q}}$. If C is not contained in a special subvariety of positive codimension, then $C(\mathbb{C}) \cap \mathcal{S}^{\mathrm{nss},[2]}$ is finite.*

Variants of Pink's conjecture for semi-abelian varieties are Zilber's conjecture 2 [Zil02] and Pink's conjecture 5.1 [Pin05]. In fact, the general version of Pink's conjecture implies the latter. Several results are known in the semi-abelian setting; we discuss two. For the moment, let C be an irreducible curve defined over $\mathbb{C}$ and embedded in the algebraic torus $\mathbb{G}_m^n$. Let us assume that C is not contained in a proper algebraic subgroup of $\mathbb{G}_m^n$. Bombieri *et al.* [BMZ08] showed

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

that $C(\mathbb{C})$ contains only finitely many points in an algebraic subgroup of codimension at least two. Their result relies on the important case, proved by Maurin [Mau08], where the curve is defined over $\overline{\mathbb{Q}}$. The theorem of Bombieri *et al.* can be seen as an analogue of the conjecture stated above with $Y(1)^n$ replaced by the algebraic torus.

The Mordell–Lang conjecture governs the intersection of subvarieties of semi-abelian varieties with the division closure of a finitely generated subgroup. This conjecture is a theorem due to work of Hindry, Faltings, McQuillan, Vojta and others. Now $Y(1)^n$ lacks a group structure having a meaningful connection to moduli problems. So we should not look at (the division closure of) finitely generated subgroups of $Y(1)^n$. Instead we have Hecke orbits. We may regard the following theorem as a variant of the Mordell–Lang conjecture for $Y(1)^n$.

We call an irreducible subvariety of $Y(1)^n$ *geodesic* if it is an arbitrary product of special subvarieties, in smaller $Y(1)^{n'}$, and singletons. A geodesic subvariety is special if and only if it contains a special point.

Let U be a subset of $\{1,\dots,n\}\times Y(1)(\mathbb{C})$. A point $(x_1,\dots,x_n)\in Y(1)(\mathbb{C})^n$ will be called *U-special* if, for each i, either x_i is special or there exists $(i,u)\in U$ such that x_i is in the Hecke orbit of u; we refer again to § 2.1 for the definition of Hecke orbits. A geodesic subvariety of $Y(1)^n$ will be called *U-special* if it contains at least one U-special point or, equivalently, if it contains a Zariski dense set of U-special points.

Theorem 3. *Let $V\subset Y(1)^n$ be a subvariety defined over $\mathbb{C}$ and $U\subset\{1,\dots,n\}\times Y(1)(\overline{\mathbb{Q}})$ a finite subset. Then V contains only finitely many maximal U-special subvarieties.*

In this theorem, there is no restriction on the dimension of the subvariety V or its field of definition.

The paper is organized in the following way. Section 2 contains much of the notation used throughout the paper. In particular, we introduce height functions and recall the notion of an o-minimal structure. A modular variant of Ax's theorem mentioned above is proved in § 3. Section 4 contains the lower bounds for Galois orbits. Section 5 synthesizes the preceding work using arguments from o-minimality into a proof of Theorems 1 and 2. Finally, Theorem 3 is proved in § 6.

2. Preliminaries

2.1 Special subvarieties

In this subsection we describe the special subvarieties of $Y(1)^n$.

The group $\mathrm{GL}_2(\mathbb{R})^+$ of real 2×2-matrices with positive determinant acts on $\mathbb{H}$ by fractional linear transformations. The modular j-invariant is an $\mathrm{SL}_2(\mathbb{Z})$-invariant, holomorphic function $j:\mathbb{H}\to\mathbb{C}=Y(1)(\mathbb{C})$. By abuse of notation we will denote by j also the Cartesian product of this function mapping $\mathbb{H}^n\to\mathbb{C}^n=Y(1)^n(\mathbb{C})$.

We define special subvarieties, following Edixhoven [Edi05], as well as geodesic subvarieties, in $Y(1)^n$ and $\mathbb{H}^n$. The word 'geodesic' is borrowed from the terminology 'totally geodesic' in Moonen [Moo98]. It is convenient to use the same terminology for the subvarieties in $Y(1)^n$ as for the corresponding subvarieties in $\mathbb{H}^n$. Thus the image under j of a geodesic (respectively special) subvariety in $\mathbb{H}^n$ will be a geodesic (respectively special) subvariety in $Y(1)^n$. An irreducible component of the inverse image under j of a geodesic (respectively special) subvariety of $Y(1)^n$ will be a geodesic (respectively special) subvariety of $\mathbb{H}^n$.

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

A complex algebraic hypersurface, or just hypersurface, of $\mathbb{H}^n$ will mean a non-empty set of the form $Y(\mathbb{C}) \cap \mathbb{H}^n$ where Y is the set of zeros of a non-zero polynomial in complex coefficients and n-variables. As $\mathbb{H}^n$ is open in $\mathbb{C}^n$ with respect to the Euclidean topology, such a hypersurface will be Zariski dense in Y.

We now come to the definition of special points.

(i) A *special point* in $\mathbb{H}^n$ is a point $(z_1, \ldots, z_n) \in \mathbb{H}^n$ with $[\mathbb{Q}(z_i) : \mathbb{Q}] = 2$ for $i = 1, \ldots, n$.

(ii) A *special point* in $Y(1)^n$ is the image under j of a special point in $\mathbb{H}^n$.

Equivalently, a special point in $Y(1)$ is the j-invariant of an elliptic curve over $\mathbb{C}$ with complex multiplication, and a special point in $Y(1)^n$ is an n-tuple of such j-invariants. It is a classical fact that special points are algebraic integers, see [Lan87, Theorem 4, p. 57].

Write $\mathbb{H}_\ell$ for the ℓth factor of $\mathbb{H}^n$. A *geodesic subvariety* of $\mathbb{H}^n$ is a subvariety $Y \subset \mathbb{H}^n$ for which there is a partition $(S_0, \ldots, S_r)$ of $\{z_1, \ldots, z_n\}$, in which only S_0 is permitted to be empty and $r = 0$ is permitted, such that Y is the Cartesian product $Y = \prod_{i=0}^{r} Y_i$ of subvarieties $Y_i \subset \prod_{\ell \in S_i} \mathbb{H}_\ell$ with the Y_i of one of the following forms.

(i) The set Y_0 is a single point in $\prod_{\ell \in S_0} \mathbb{H}_\ell$.

(ii) For $i > 0$, the variety Y_i is the image of $\mathbb{H}$ under a map $z \mapsto (g_\ell z) \in \prod_{\ell \in S_i} \mathbb{H}_\ell$ with each $g_\ell \in \mathrm{GL}_2(\mathbb{Q})^+$, the group of matrices in $\mathrm{GL}_2(\mathbb{R})^+$ with rational entries.

A *geodesic subvariety* of $Y(1)^n$ is the image under j of a geodesic subvariety of $\mathbb{H}^n$.

We let $\Phi_N \in \mathbb{Z}[X, Y]$ denote the modular polynomial of order N; for a definition and properties we refer to [Lan87, ch. 5]. Observe that if a relation $z_j = g z_j$ holds for coordinates z_i, z_j on $\mathbb{H}^n$, with $g \in \mathrm{GL}_2(\mathbb{Q})^+$, then $\Phi_N(x_i, x_j)$ holds, for a suitable N, on the projection. By [Edi05, Proposition 3.1], a geodesic subvariety of $Y(1)^n$ may be equivalently defined as an irreducible component of the locus defined by requiring that certain coordinates be constant, and certain pairs of coordinates be related by some modular polynomial.

Special subvarieties are an important kind of geodesic subvarieties.

(i) A *special* subvariety of $\mathbb{H}^n$ is a geodesic subvariety of $\mathbb{H}^n$ for which Y_0 is empty or each coordinate of Y_0 is a special point of $\mathbb{H}$.

(ii) A *special subvariety* of $Y(1)^n$ is the image under j of a special subvariety in $\mathbb{H}^n$.

One observes the following.

(i) A special point in $\mathbb{H}^n$ is a special subvariety in $\mathbb{H}^n$ of dimension zero.

(ii) A special point of $Y(1)^n$ is a special subvariety of $Y(1)^n$ of dimension zero.

Points $z_1, z_2 \in \mathbb{H}$ will be called *Hecke equivalent* and said to be in the same *Hecke orbit* if there exists $g \in \mathrm{GL}_2(\mathbb{Q})^+$ such that $z_1 = g z_2$. Points $x_1, x_2 \in Y(1)(\mathbb{C})$ will be called *Hecke equivalent* and said to be in the same *Hecke orbit* if there exist $z_1, z_2 \in \mathbb{H}$ such that $j(z_1) = x_1$, $j(z_2) = x_2$, and z_1, z_2 are Hecke equivalent.

Let U be a finite subset of $\{1, \ldots, n\} \times \mathbb{H}$. A point $(z_1, \ldots, z_n) \in \mathbb{H}^n$ will be called *U-special* if, for each $i = 1, \ldots, n$, either z_i is special or there exists $(i, u) \in U$ such that z_i is in the Hecke orbit of u. A *U-special subvariety* of $\mathbb{H}^n$ is a U-special point or a geodesic subvariety of $\mathbb{H}^n$ of positive dimension that contains a u-special point. For such U write $j(U) = \{(i, j(u)); (i, u) \in U\}$. The image under j of a U-special point in $\mathbb{H}^n$ is then a $j(U)$-special point of $Y(1)^n$. The image under j of a U-special subvariety of $\mathbb{H}^n$ is likewise a $j(U)$-special subvariety of $Y(1)^n$.

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

2.2 Heights

Later we will recall results describing the distribution of rational points on definable sets. In order to formulate these we need to define suitable height functions.

The *absolute logarithmic height* $h(x)$ of an algebraic number x is defined as follows. Let $P = a_d X^d + \cdots + a_0$ be a polynomial with integer coefficients and irreducible as an element of $\mathbb{Z}[X]$ with $a_d \neq 0$ and $P(x) = 0$. Then P is defined uniquely up to sign and we set

$$h(x) = \frac{1}{d} \log\left(|a_d| \prod_{P(x')=0} \max\{1, |x'|\} \right)$$

where the product runs over the complex roots of P. Then $h(x)$ coincides with the height of the projective point $[1 : x]$ as in [BG06, ch. 1.5], cf. also Proposition 1.6.6. The *absolute exponential height* is $H(x) = \exp h(x)$.

For example, if p and q are coprime integers with $q \neq 0$, then $H(p/q) = \max\{|p|, |q|\}$.

If $x = (x_1, \ldots, x_n) \in \overline{\mathbb{Q}}^n$, we define

$$h(x) = \max\{h(x_1), \ldots, h(x_n)\} \quad \text{and} \quad H(x) = \max\{H(x_1), \ldots, H(x_n)\}.$$

Later on we need the height of matrices with algebraic coefficients and we obtain this by identifying a matrix with a point in a suitable power of $\overline{\mathbb{Q}}$.

We simply call any of these various height functions the *height*.

2.3 On o-minimal structures, definable sets, and rational points

For an introduction to o-minimal structures see the book [Dri98] by van den Dries. The notion originated in work of van den Dries [Dri84] and Pillay–Steinhorn [PS86]. Here we give the briefest sketch and the formal definitions required to state the theorems we use.

A structure over $\mathbb{R}$ is a collection $\mathfrak{S}$ of sets in $\mathbb{R}^n$, $n = 1, 2, \ldots$ with closure properties corresponding to definability in a suitable first order language. In particular $\mathfrak{S}$ should contain every semi-algebraic set defined over $\mathbb{R}$ and be closed under Cartesian products, Boolean operations (i.e. finite union, finite intersection, and complement), and coordinate projections. This makes the structure rich enough to allow many constructions. It is *o-minimal* ('order-minimal') if, notwithstanding these closure properties, the subsets of $\mathbb{R}$ that belong to $\mathfrak{S}$ are all finite unions of points and (possibly unbounded) intervals. The formal definition is as follows.

A *pre-structure* is a sequence $\mathfrak{S} = (\mathfrak{S}_\nu : \nu \geqslant 1)$ where each $\mathfrak{S}_\nu$ is a collection of subsets of $\mathbb{R}^\nu$. A pre-structure $\mathfrak{S}$ is called a *structure over* $\mathbb{R}$ if, for all $\nu, \mu \geqslant 1$, the following conditions are satisfied.

(i) The set $\mathfrak{S}_\nu$ is a Boolean algebra under the usual set-theoretic operations.

(ii) The set $\mathfrak{S}_\nu$ contains every semi-algebraic subset of $\mathbb{R}^\nu$.

(iii) If $A \in \mathfrak{S}_\nu$ and $B \in \mathfrak{S}_\mu$ then $A \times B \in \mathfrak{S}_{\nu+\mu}$.

(iv) If $\mu \geqslant \nu$ and $A \in \mathfrak{S}_\mu$ then $\pi(A) \in \mathfrak{S}_\nu$, where $\pi : \mathbb{R}^\mu \to \mathbb{R}^\nu$ is projection onto the first ν coordinates.

If, in addition, the following condition holds then $\mathfrak{S}$ is called an *o-minimal* structure over $\mathbb{R}$.

(v) The boundary of every set in $\mathfrak{S}_1$ is finite.

For brevity we usually drop the addition 'over $\mathbb{R}$' when referring to structures or o-minimal structures. If $\mathfrak{S}$ is a structure, and $Z \in \mathfrak{S}_\nu$, we say Z is *definable in* $\mathfrak{S}$. Let $A \subset \mathbb{R}^\nu$ and $B \subset \mathbb{R}^\mu$ be

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

subsets. A function $f : A \to B$ is said to be *definable in* $\mathfrak{S}$ if its graph $\{(a, b) \in A \times B; b = f(a)\}$ is definable in $\mathfrak{S}$. If $f : A \to B$ is definable in $\mathfrak{S}$ then A and the image of f are definable in $\mathfrak{S}$. By a *definable family* in $\mathfrak{S}$ we mean a set $Z \subset \mathbb{R}^\nu \times \mathbb{R}^\mu$ definable in $\mathfrak{S}$, considered as the family of fibers

$$Z_y = \{x \in \mathbb{R}^\nu; (x, y) \in Z\}, \quad y \in \mathbb{R}^\mu.$$

For brevity we will often omit reference to the structure when speaking of definable sets, functions, families, etc.

If Z is a definable family then the set $Y = \{y \in \mathbb{R}^\mu; Z_y \neq \emptyset\}$ is definable, so it will be immaterial whether we consider quantifications over Y or $\mathbb{R}^\mu$. Note that we consider the fiber Z_y to be a subset of $\mathbb{R}^\nu$, so any rationality considerations relate to the $\mathbb{R}^\nu$-coordinates and not to the coordinates of the parameter $y \in \mathbb{R}^\mu$.

For our purposes it suffices to consider sets definable in the o-minimal structure $\mathbb{R}_{\mathrm{an,exp}}$. The 'exp' indicates that this structure contains the graph of exponential function $\exp : \mathbb{R} \to \mathbb{R}$ while the 'an' indicates that it contains the graphs of all restricted analytic functions $f|_{[0,1]^n} : [0, 1]^n \to \mathbb{R}$ where f is real analytic in a neighborhood of $[0, 1]^n$. The o-minimality of $\mathbb{R}_{\mathrm{exp}}$, the smallest structure containing the graph of exp, is due to Wilkie [Wil96]. That of $\mathbb{R}_{\mathrm{an}}$, the structure generated by restricted analytic functions, follows from Gabrielov's theorem [Gab68] as observed by van den Dries [Dri86]. The o-minimality of the structure $\mathbb{R}_{\mathrm{an,exp}}$ generated by their union is due to van den Dries and Miller [DM94], see also [DMM94].

Let $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$ denote real and imaginary part of a complex number z. Using $\mathrm{Re}(z), \mathrm{Im}(z)$ we may identify subsets of $\mathbb{C}^n$ with subsets of $\mathbb{R}^{2n}$. The set

$$F = \{z \in \mathbb{H}; |z| \geqslant 1, \mathrm{Re}(z) \in [-1/2, 1/2)\} \setminus \{z \in \mathbb{H}; |z| = 1, \mathrm{Re}(z) \in (0, 1/2)\}$$

is a *fundamental domain* for the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{H}$, in other words it meets any orbit in precisely one point. Observe that F is a semi-algebraic subset of $\mathbb{R}^2$ under the above identification. The Cartesian product F^n is then a fundamental domain for the component-wise action of $\mathrm{SL}_2(\mathbb{Z})^n$ on $\mathbb{H}^n$.

Theorem 4. *The restriction $j|_F : F \to \mathbb{C}$ is definable in $\mathbb{R}_{\mathrm{an,exp}}$.*

This follows from a result of Peterzil and Starchenko [PS04] on definability of the Weierstrass $\wp$-function as a function of two variables, but is alternatively deduced from the q-expansion of the j-function.

Hence $j|_{F^n} : F^n \to \mathbb{C}^n$ is a definable function and $Z = j^{-1}(C) \cap F$ is a definable set. Indeed, all the definability properties we need follow from Theorem 2.1 combined with standard properties of o-minimal structures (e.g. as set out in [Dri98, DM96]). So, rather than working in $\mathbb{R}_{\mathrm{an,exp}}$ we could work in the (much smaller) o-minimal structure generated by the graph of j restricted to F as a set in $\mathbb{R}^2 \times \mathbb{R}^2$.

A result of Wilkie and the second author [PW06] concerns rational points up to a given height on a set X definable in any o-minimal structure. The result asserts that there are 'few' such points that do not lie on some connected positive dimensional semi-algebraic set contained in X. The result we will use here is a refinement articulated in [Pil11].

For a positive integer k set

$$X(k, T) = \{(x_1, \ldots, x_n) \in X; [\mathbb{Q}(x_i) : \mathbb{Q}] \leqslant k \text{ and } H(x_i) \leqslant T \text{ for } i = 1, \ldots, n\}.$$

A *definable semi-algebraic block* or *block* of dimension w in $\mathbb{R}^n$ is a connected definable set $X \subset \mathbb{R}^n$ for which there is a semi-algebraic set $A \subset \mathbb{R}^n$ such that every point in X has a neighborhood which coincides with a smooth neighborhood of A of dimension w.

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

A *definable semi-algebraic block family* or *block family* of dimension w in $\mathbb{R}^n$ is a definable family $W \subset \mathbb{R}^n \times \mathbb{R}^m$ such that for each $y \in \mathbb{R}^m$ the fiber W_y is either empty or a block of dimension w. For brevity we sometimes omit the reference to the dimension w below.

For example, a block of dimension zero is a point; the set $\{(x, y) \in \mathbb{R}^2; 0 < y < e^x\}$ is a block of dimension two and degree one in $\mathbb{R}_{\exp}$; the collection $X_{(a,b)} = \{(x, y) \in \mathbb{R}^2; a < x < b, 0 < y < e^x\}$, where $a, b \in \mathbb{R}$, $a < b$ is a block family.

With these definitions we can state the form of the result we require. It shows that, given a definable set Z and $\epsilon > 0$, the algebraic points with a fixed bounded degree d in Z up to height T are contained in $c(Z, d, \epsilon)T^\epsilon$ blocks contained in Z that come from a finite number of block families; here $c(Z, d, \epsilon) > 0$ is independent of T. This result is even uniform over definable families, i.e. we have the following result.

Theorem 5 [Pil11, Theorem 3.6]. *Let $X \subset \mathbb{R}^n \times \mathbb{R}^m$ be a definable family in some o-minimal structure over $\mathbb{R}$ with fibers X_y, $y \in \mathbb{R}^m$. Let $\epsilon > 0$ and $k \geqslant 1$. There is a finite number $J(X, k, \epsilon)$ of block families,*

$$W^{(j)} \subset \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{\ell_j}, \quad j = 1, \dots, J(X, k, \epsilon),$$

and a constant $c(X, k, \epsilon)$ with the following properties.

(i) *For all $(y, z) \in \mathbb{R}^m \times \mathbb{R}^{\ell_j}$ we have $W_{(y,z)} \subset X_y$.*

(ii) *For all $y \in \mathbb{R}^m$ and $T \geqslant 1$ the set $X_y(k, T)$ is contained in the union of at most*

$$c(X, k, \epsilon)T^\epsilon$$

blocks of the form $W^{(j)}_{(y,z)}$ for suitable $j = 1, \dots, J(X, k, \epsilon)$ and $z \in \mathbb{R}^{\ell_j}$.

3. Algebraic independence of modular logarithms

In this section we prove a weak version of Ax's theorem [Ax71] for the modular j-function. The content of this section's main result, Proposition 3.1, can be described as follows. If the restriction of a local inverse of $j : \mathbb{H}^n \to \mathbb{C}^n$ to an irreducible curve in $Y(1)^n$ takes values in a complex algebraic hypersurface of $\mathbb{H}^n$, then the curve is contained in a geodesic subvariety of positive codimension. The proof uses a Hodge theoretic result of André [And92]. Daniel Bertrand has kindly suggested an alternative approach to the following proposition based on [Del71, Deligne's corollaire 4.4.13].

Proposition 3.1. *Let $C \subset Y(1)^n$ be an irreducible curve defined over $\mathbb{C}$ and let $z \in j^{-1}(C(\mathbb{C}))$. We assume that there exists a complex algebraic hypersurface of $\mathbb{H}^n$ that contains a neighborhood of z in $j^{-1}(C(\mathbb{C})) \subset \mathbb{H}^n$. Then C is contained in a geodesic subvariety of positive codimension.*

For technical reasons much of the proof deals with curves in $Y(2)^n$ where $Y(2)$ is the modular curve $\mathbb{P}^1 \smallsetminus \{0, 1, \infty\}$ with level 2 structure. We will return to $Y(1)$ by using the morphism $Y(2)^n \to Y(1)^n$ which is defined component-wise by

$$Y(2)(\mathbb{C}) \ni \lambda \mapsto 2^8 \frac{(\lambda^2 - \lambda + 1)^3}{\lambda^2(\lambda - 1)^2}. \tag{1}$$

Any such λ leads to an elliptic curve in Legendre form with parameter λ, i.e. the elliptic curve cut out by the cubic

$$y^2 z - x(x - z)(x - \lambda z) \tag{2}$$

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

in $\mathbb{P}^2$. The right-hand side of (1) is its j-invariant by [Hus04, Remark 1.4, p. 87]. If we regard λ as an independent, then the polynomial (2) determines a subvariety $\mathcal{E} \subset \mathbb{P}^2 \times Y(2)$. Then $\mathcal{E}$ is an abelian scheme over $Y(2)$ where the structural morphism is the projection on the parameter. For $\lambda \in Y(2)(\mathbb{C})$ we set $\mathcal{E}_\lambda$ to be the fiber of $\mathcal{E} \to Y(2)$ above λ. This is just the elliptic curve cut out by (2).

A period lattice of an abelian variety A defined over $\mathbb{C}$ is a discrete subgroup $\Omega \subset \mathbb{C}^{\dim A}$ of rank two $\dim A$ such that $A(\mathbb{C})$ and $\mathbb{C}^{\dim A}/\Omega$ are isomorphic complex tori. The period lattice is determined uniquely up to homothety if A is a fiber of $\mathcal{E} \to Y(2)$. We proceed by choosing such a lattice in this case.

Gauss's hypergeometric function

$$ {}_2F_1\left(\frac{1}{2}, \frac{1}{2}, 1, \lambda\right) = \sum_{n=0}^{\infty} \frac{(2n)!^2}{2^{4n} n!^4} \lambda^n $$

converges for $\lambda \in \mathbb{C}$ with $|\lambda| < 1$. We use two holomorphic functions

$$ \omega_1(\lambda) = {}_2F_1\left(\frac{1}{2}, \frac{1}{2}, 1, \lambda\right)\pi \quad \text{and} \quad \omega_2(\lambda) = {}_2F_1\left(\frac{1}{2}, \frac{1}{2}, 1, 1-\lambda\right)\pi i; \tag{3} $$

they are both defined on

$$ \{\lambda \in \mathbb{C};\ |\lambda| < 1 \text{ and } |1-\lambda| < 1\}. \tag{4} $$

A loop l in any topological space represents an element $[l]$ of the corresponding fundamental group. The fundamental group $\pi_1(Y(2)(\mathbb{C}), 1/2)$ is a free group of rank two. Let l_0 and l_1 be loops based at $1/2$ which circle around 0 and 1, respectively, precisely once counterclockwise without circling around 1 and 0, respectively. Then $[l_0]$ and $[l_1]$ generate $\pi_1(Y(2)(\mathbb{C}), 1/2)$. For all $\lambda \in Y(2)(\mathbb{C})$ we fix once and for all a path leading from $1/2$ to λ. These choices enable us to identify $\pi_1(Y(2)(\mathbb{C}), 1/2)$ with $\pi_1(Y(2)(\mathbb{C}), \lambda)$; hence we consider $[l_{0,1}]$ as elements of any $\pi_1(Y(2)(\mathbb{C}), \lambda)$. We define a group homomorphism $\rho : \pi_1(Y(2)(\mathbb{C}), \lambda) \to \mathrm{SL}_2(\mathbb{Z})$ by setting

$$ \rho([l_i]) = \gamma_i \quad \text{with } \gamma_0 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \gamma_1 = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}. \tag{5} $$

These matrices generate a subgroup Γ_2 of $\mathrm{SL}_2(\mathbb{Z})$. One can show that Γ_2 has index 12 in $\mathrm{SL}_2(\mathbb{Z})$. However, we will only need the weaker fact that Γ_2 is infinite; indeed, neither matrix in (5) has finite order.

The next statement is classical.

Lemma 3.1. (i) *If λ is in (4), then $(\omega_2(\lambda), \omega_1(\lambda))$ is a basis of a period lattice of $\mathcal{E}_\lambda$. In particular, $\omega_{1,2}$ are $\mathbb{R}$-linearly independent.*

(ii) *Both $\omega_{1,2}$ have analytic continuations along any path in $Y(2)(\mathbb{C})$ with values constituting a basis of a period lattice of the corresponding $\mathcal{E}_\lambda$. Moreover, ω_2/ω_1 has an analytic continuation along any path in $Y(2)(\mathbb{C})$ which takes values in $\mathbb{H}$. Let $\lambda \in Y(2)(\mathbb{C})$; we continue $\omega_{1,2}$ along the chosen path from $1/2$ to λ and obtain a basis $(\widetilde{\omega}_2, \widetilde{\omega}_1)$ of a period lattice of $\mathcal{E}_\lambda$. Continuing $\omega_{1,2}$ further along any loop l in $Y(2)(\mathbb{C})$ based at λ gives an action on this period lattice represented by $\rho([l])$ with respect to $(\widetilde{\omega}_2, \widetilde{\omega}_1)$.*

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

(iii) *There is a holomorphic map* $u : \mathbb{H} \to Y(2)(\mathbb{C})$ *such that*

$$\begin{array}{ccc} \mathbb{H} & \xrightarrow{u} & Y(2)(\mathbb{C}) \\ {\scriptstyle j}\downarrow & & \downarrow \\ \mathbb{C} & = & Y(1)(\mathbb{C}) \end{array} \tag{6}$$

commutes; here the vertical arrow on the right is (1). An analytic continuation of ω_2/ω_1 *along any path in* $Y(2)(\mathbb{C})$ *is locally a right-inverse of* u.

Proof. Part (i) follows from the discussion in [Hus04, ch. 9.6].

The fact that $\omega_{1,2}$ have analytic continuation in part (ii) follows from [EMOT81, ch. 2.1]. We remark that $\omega_2(1/2)/\omega_1(1/2) = i \in \mathbb{H}$ with $\omega_{1,2}$ as in (3). The same references imply the fact that such continuations again lead to period lattices and that ω_2/ω_1 takes values in $\mathbb{H}$. It suffices to show the statement on the action of $\pi_1(Y(2)(\mathbb{C}), \lambda)$ for the loops l_0 and l_1 based at $\lambda = 1/2$; this is done in [EMOT81, ch. 2.7.1].

The third part is a consequence of the following observation; for details we refer to [Hus04, ch. 9]. For $z \in \mathbb{H}$ let $\wp(\cdot; z)$ be the Weierstrass function attached to the lattice $\mathbb{Z} + \mathbb{Z}z$. Then $e_1(z) = \wp(z/2; z)$, $e_2(z) = \wp(1/2; z)$, and $e_3(z) = \wp((1+z)/2; z)$ are holomorphic functions on $\mathbb{H}$. For fixed z, the Weierstrass function and its derivative satisfy the differential equation

$$\wp'(\cdot; z)^2 = 4\wp(\cdot; z)^3 - g_2(z)\wp(\cdot; z) - g_3(z) \tag{7}$$

for some $g_{2,3}(z) \in \mathbb{C}$. This equation determines an elliptic curve in Weierstrass form. Moreover, $e_{1,2,3}(z)$ are the three distinct roots of $4T^3 - g_2(z)T - g_3(z)$. Hence we can define the holomorphic function $u(z) = (e_3(z) - e_1(z))/(e_2(z) - e_1(z)) \in \mathbb{C} \smallsetminus \{0, 1\} = Y(2)(\mathbb{C})$ on $\mathbb{H}$. Let us set $T(\lambda) = \omega_2(\lambda)/\omega_1(\lambda)$ with λ in (4). Then $\mathcal{E}_{u(T(\lambda))}(\mathbb{C})$ is isomorphic to the complex torus $\mathbb{C}/(\mathbb{Z} + \mathbb{Z}T(\lambda))$. Hence the right-hand side of (1) with λ replaced by $u(T(\lambda))$ equals $j(T(\lambda))$. This shows that (6) commutes on the open set $\{T(\lambda); \lambda \text{ in } (4)\} \subset \mathbb{H}$. By analyticity the diagram commutes on all of $\mathbb{H}$. However, $\mathbb{C}/(\mathbb{Z} + \mathbb{Z}T(\lambda))$ is isomorphic to $\mathcal{E}_\lambda(\mathbb{C})$. Comparing j-invariants and a small calculation show $u(T(\lambda)) = \lambda$. This identity must hold along any analytic continuation of T. □

Let $C \subset Y(2)^n$ be an irreducible curve defined over $\mathbb{C}$.

The set C^{ns} of smooth complex points of C carries the structure of a complex manifold. By abuse of notation we sometimes also consider C^{ns} as an irreducible algebraic curve. We let $x = (\lambda_1, \dots, \lambda_n) \in C^{\mathrm{ns}}$ denote a base point, to be chosen later on. Each λ_i determines a basis $(\widetilde{\omega}_{i2}, \widetilde{\omega}_{i1})$ of a period lattice of $\mathcal{E}_{\lambda_i}$ as in Lemma 3.1(ii) using the fixed path from $1/2$ to λ_i. We set $z = (\widetilde{\omega}_{1,2}/\widetilde{\omega}_{1,1}, \dots, \widetilde{\omega}_{n2}/\widetilde{\omega}_{n1}) \in \mathbb{H}^n$. By abuse of notation we use u to denote the n-fold power $\mathbb{H}^n \to Y(2)(\mathbb{C})^n$. Then $u(z) = x$ by Lemma 3.1(iii).

We identify $\pi_1(Y(2)^n(\mathbb{C}), x)$ with $\prod_{i=1}^n \pi_1(Y(2)(\mathbb{C}), \lambda_i)$ in the natural way. Using φ we obtain a homomorphism $\pi_1(Y(2)^n(\mathbb{C}), x) \to \mathrm{SL}_2(\mathbb{Z})^n$. Finally, the inclusion $C^{\mathrm{ns}} \subset Y(2)^n(\mathbb{C})$ induces a homomorphism $\pi_1(C^{\mathrm{ns}}, x) \to \pi_1(Y(2)^n(\mathbb{C}), x)$. We call the composition of these two homomorphisms $\rho : \pi_1(C^{\mathrm{ns}}, x) \to \mathrm{SL}_2(\mathbb{Z})^n$.

The n-fold product $\mathcal{E}^n \to Y(2)^n$ is also an abelian scheme. We pull it back using the inclusion morphism $C^{\mathrm{ns}} \hookrightarrow Y(2)^n$ and obtain a new abelian scheme $\mathcal{A} \to C^{\mathrm{ns}}$. The homomorphism ρ describes the action of $\pi_1(C^{\mathrm{ns}}; x)$ on the period lattice of the fiber $\mathcal{A}_x$ of $\mathcal{A} \to C^{\mathrm{ns}}$ above x in terms of the basis $(\widetilde{\omega}_{1,2}v_1, \widetilde{\omega}_{1,1}v_1, \dots, \widetilde{\omega}_{n2}v_n, \widetilde{\omega}_{n1}v_n)$; here $(v_1, \dots, v_n)$ is the standard basis of $\mathbb{C}^n$. Let us define $\Gamma = \rho(\pi_1(C^{\mathrm{ns}}, x))$. Then Γ is a subgroup of $\Gamma_2^n \subset \mathrm{SL}_2(\mathbb{Z})^n$.

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

There is a natural identification of the period lattice of $\mathcal{A}_x$ with the homology group $H_1(\mathcal{A}_x(\mathbb{C}), \mathbb{Z})$. The action of $\pi_1(C^{\mathrm{ns}}, x)$ on $H_1(\mathcal{A}_x(\mathbb{C}), \mathbb{Z})$ is central in Hodge theory and can also be introduced coordinate-free. We have chosen to work with a fixed basis to make some calculations more explicit.

The group $\mathrm{SL}_2(\mathbb{Z})^n$ acts on $\mathbb{H}^n$ by component-wise fractional linear transformations.

Lemma 3.2. *Say there is a neighborhood of z in $u^{-1}(C(\mathbb{C}))$ that is contained in a complex algebraic hypersurface Σ of $\mathbb{H}^n$. Then $\Gamma^{\mathsf{T}} z \subset \Sigma$ where $^{\mathsf{T}}$ means taking the transpose of all coordinates of all elements.*

Proof. Let $[l] \in \pi_1(C^{\mathrm{ns}}, x)$ be represented by a loop $l : [0, 1] \to C^{\mathrm{ns}}$. We continue $\lambda \mapsto \omega_2(\lambda)/\omega_1(\lambda)$ analytically along each coordinate of l as in Lemma 3.1(ii). This leads to a new path $[0, 1] \to \mathbb{H}^n$ starting at z. Each image point lies in $u^{-1}(C^{\mathrm{ns}})$ by Lemma 3.1(iii). By analyticity and because a neighborhood of z in $u^{-1}(C(\mathbb{C}))$ is contained in Σ we conclude that the end point of the new path lies in Σ too. If

$$\rho([l]) = \left(\begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}, \ldots, \begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix} \right)$$

then the bases $(\widetilde{\omega}_{i2}, \widetilde{\omega}_{i1})$ transform to

$$\begin{bmatrix} \widetilde{\omega}_{i2} & \widetilde{\omega}_{i1} \end{bmatrix} \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix}$$

and the quotients $\widetilde{\omega}_{i2}/\widetilde{\omega}_{i1}$ transform to

$$\begin{bmatrix} a_i & c_i \\ b_i & d_i \end{bmatrix} \frac{\widetilde{\omega}_{i2}}{\widetilde{\omega}_{i1}} \in \mathbb{H}.$$

However, these are the coordinates of $\rho([l])^{\mathsf{T}} z$. We conclude the lemma since $[l]$ was arbitrary. □

We consider SL_n and GL_n as algebraic groups over $\mathbb{Q}$. For $i = 1, \ldots, n$ we let $p_i : \mathrm{SL}_2^n \to \mathrm{SL}_2$ denote the projection onto the ith factor. The following lemma describes certain normal algebraic subgroups of SL_2^n.

Lemma 3.3. *Let G be a normal algebraic subgroup of SL_2^n. If $p_i(G) = \mathrm{SL}_2$ for $i = 1, \ldots, n$, then $G = \mathrm{SL}_2^n$.*

Proof. Let us assume that the conclusion is false. By a result of Kolchin [Kol68] there is $\alpha \in \mathrm{SL}_2(\mathbb{C})$ such that

$$H(\mathbb{C}) = \{(g, \chi(g)\alpha g \alpha^{-1}); g \in \mathrm{SL}_2(\mathbb{C})\} \quad \text{with } \chi(g) \in \{\pm 1\}$$

where H is the projection of G onto two appropriate and distinct factors of SL_2^n. Now H is normal in SL_2^2 since G is normal in SL_2^n. For each triple $g, g', g'' \in \mathrm{SL}_2(\mathbb{C})$ we have

$$(g', g'')(g, \chi(g)\alpha g \alpha^{-1})(g', g'')^{-1} \in H(\mathbb{C}),$$

or in other words

$$\alpha g' g g'^{-1} \alpha^{-1} = g'' \alpha g \alpha^{-1} g''^{-1}.$$

This leads to a contradiction as follows. We first specialize $g = g'$ to see

$$\alpha g \alpha^{-1} = g''(\alpha g \alpha^{-1}) g''^{-1}$$

for all $g, g'' \in \mathrm{SL}_2(\mathbb{C})$. Hence $\alpha g \alpha^{-1}$ is in $\{\pm 1\}$, the center of $\mathrm{SL}_2(\mathbb{C})$, for all $g \in \mathrm{SL}_2(\mathbb{C})$. This absurdity establishes our lemma. □

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

We regard the set of complex points of any variety defined over $\mathbb{C}$ as a complex analytic space [GR84]. In this article, all complex algebraic spaces are reduced.

Lemma 3.4. *Let us assume that there is a neighborhood of a smooth point of dimension one of the complex analytic space $u^{-1}(C(\mathbb{C}))$ that is contained in a complex algebraic hypersurface of $\mathbb{H}^n$. Then C is contained in a geodesic subvariety of positive codimension.*

Proof. We may assume that no coordinate function is constant when restricted to C.

Two elliptic curves over $\mathbb{C}$ are isogenous if and only if their j-invariants are coordinates of a complex point of $Y(1)^2$ where some modular polynomial Φ_N vanishes. Say $1 \leqslant i < i' \leqslant n$. We recall that $x = (\lambda_1, \dots, \lambda_n) \in C^{\mathrm{ns}}$ is our base point which we have yet to fix. By considering the right-hand side of (1) we see that $\mathcal{E}_{\lambda_i}$ and $\mathcal{E}_{\lambda_{i'}}$ are isogenous if and only if

$$\Phi_N\left(2^8\frac{(\lambda_i^2-\lambda_i+1)}{\lambda_i^2(\lambda_i-1)^2}, 2^8\frac{(\lambda_{i'}^2-\lambda_{i'}+1)}{\lambda_{i'}^2(\lambda_{i'}-1)^2}\right) = 0 \tag{8}$$

for some positive integer N.

If such a relation holds identically on C^{ns} for some N, then it holds identically on C and we are done. Therefore, let us assume the contrary; we explain now how this leads to a contradiction.

A single relation (8) holds for only finitely many points on C. Hence there are at most countably many points on $C(\mathbb{C})$ for which there exist a positive integer N and $1 \leqslant i < i' \leqslant n$ with (8). No coordinate function restricted to C is constant so there are at most countably many points on $C(\mathbb{C})$ having some coordinate correspond to an elliptic curve with complex multiplication.

Let z be a smooth point of the complex analytic space $u^{-1}(C(\mathbb{C}))$ with $\dim_z u^{-1}(C(\mathbb{C})) = 1$. We may assume that $u(z) \in C^{\mathrm{ns}}$ and even that $u(z)$ lies outside a prescribed countable subset of C^{ns}. Without loss of generality, no two coordinates of $u(z)$ correspond to isogenous elliptic curves and no coordinate of $u(z)$ is special. We fix $x = u(z)$.

We consider the restriction of the ith projection $Y(2)^n \to Y(2)$ to C^{ns}. After replacing C^{ns} and $Y(2)$ by sufficiently small Zariski open and non-empty subsets we obtain a finite unramified covering of Riemann surfaces. Thus, the image of the induced homomorphism $\pi_1(C^{\mathrm{ns}}, x) \to \pi_1(Y(2)(\mathbb{C}), x_i)$ has finite index in $\pi_1(Y(2)(\mathbb{C}), x_i)$. If we map the image of $\pi_1(C^{\mathrm{ns}}; x)$ using the surjective homomorphism $\varphi : \pi_1(Y(2)(\mathbb{C}), \lambda_i) \to \Gamma_2$ we get $p_i(\Gamma)$. Then $p_i(\Gamma)$ has finite index in Γ_2. We let $\overline{\Gamma}$ denote the Zariski closure of Γ in SL_2^n. This is an algebraic group with unit component $\overline{\Gamma}^0$. Now $\overline{\Gamma}^0$ has finite index in $\overline{\Gamma}$, so $\Gamma^0 = \Gamma \cap \overline{\Gamma}^0$ has finite index in Γ. It follows that $p_i(\Gamma^0)$ has finite index in Γ_2 and is therefore an infinite group. We conclude that $p_i(\overline{\Gamma}^0)$ is an algebraic group with

$$\dim p_i(\overline{\Gamma}^0) \geqslant 1. \tag{9}$$

Further down we will even show

$$\overline{\Gamma}^0 = \mathrm{SL}_2^n, \tag{10}$$

but for this we need some Hodge theory.

The fiber of $\mathcal{A} \to C^{\mathrm{ns}}$ above x is a product $A = \mathcal{E}_{\lambda_1} \times \cdots \times \mathcal{E}_{\lambda_n}$ of pairwise non-isogenous elliptic curves without complex multiplication. We recall the comments made after the proof of Lemma 3.1; there we constructed the period lattice $\Omega \subset \mathbb{C}^n$ of A together with a basis $\mathcal{B}$. Moreover, the action of $\pi_1(C^{\mathrm{ns}}, x)$ on Ω with respect to $\mathcal{B}$ is given by ρ. Now $\mathcal{B}$ is also an $\mathbb{R}$-basis of $\mathbb{C}^n$, so the $\mathbb{R}$-vector space $\Omega \otimes \mathbb{R}$ carries the structure of a $\mathbb{C}$-vector space. In particular, the group $\mathbb{C}\smallsetminus\{0\}$ acts on $\Omega \otimes \mathbb{R}$. This action and $\mathcal{B}$ define a representation $\mathbb{C}\smallsetminus\{0\} \to \mathrm{GL}_{2n}(\mathbb{R})$.

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

The smallest algebraic subgroup of GL_{2n} defined over $\mathbb{Q}$ containing the image of this representation is the *Mumford–Tate group* $\mathrm{MT}(A)$ of A. In our situation $\mathrm{MT}(A) \subset \mathrm{GL}_2^n$ where the latter is embedded diagonally in GL_{2n}. A result of Imai [Ima76] implies

$$\mathrm{SL}_2^n \subset \mathrm{MT}(A). \tag{11}$$

For an algebraic group G let G^{der} denote its derived group (G, G). André [And92, Theorem 1] states that $\overline{\Gamma}^0$ is a normal subgroup of $\mathrm{MT}(A)^{\mathrm{der}}$. Let us first determine $\mathrm{MT}(A)^{\mathrm{der}}$. Indeed, $\mathrm{MT}(A)^{\mathrm{der}}$ is a normal algebraic subgroup of $\mathrm{MT}(A)$ and $\mathrm{MT}(A)^{\mathrm{der}} \subset \mathrm{SL}_2^n$. From (11) we conclude that $\mathrm{MT}(A)^{\mathrm{der}}$ is normal in SL_2^n. We also derive $p_i(\mathrm{MT}(A)^{\mathrm{der}}) \supset p_i((\mathrm{SL}_2^n)^{\mathrm{der}}) = {\mathrm{SL}_2}^{\mathrm{der}}$ and ${\mathrm{SL}_2}^{\mathrm{der}}$ is a normal subgroup of SL_2. However, the only normal subgroups of SL_2 are $\{1\}$, $\{\pm 1\}$, and SL_2. The first two can be excluded since $\mathrm{SL}_2/{\mathrm{SL}_2}^{\mathrm{der}}$ is commutative. Hence $p_i(\mathrm{MT}(A)^{\mathrm{der}}) = \mathrm{SL}_2$ for $i = 1, \ldots, n$. Lemma 3.3 implies $\mathrm{MT}(A)^{\mathrm{der}} = \mathrm{SL}_2^n$.

André's result tells us that $\overline{\Gamma}^0$ is normal in SL_2^n and hence $p_i(\overline{\Gamma}^0)$ is normal in SL_2. From the list of normal algebraic subgroups of SL_2 and (9) we see $p_i(\overline{\Gamma}^0) = \mathrm{SL}_2$. Lemma 3.3 applies again and we conclude $\overline{\Gamma}^0 = \mathrm{SL}_2^n$. This establishes (10).

Let Σ be as in the hypothesis. By Lemma 3.2 we have $\Gamma^{\mathsf{T}} z \subset \Sigma$. We consider the embedding $\mathbb{H}^n \subset (\mathbb{P}^1)^n(\mathbb{C})$ defined component-wise by $z' \mapsto [z' : 1]$ and let $\mathrm{SL}_2(\mathbb{Z})^n$ act component-wise linearly on $(\mathbb{P}^1)^n(\mathbb{C})$ such that the restriction to $\mathbb{H}^n$ of this action is the usual one. This new action extends to an algebraic action of SL_2^n on $(\mathbb{P}^1)^n$. The Zariski closure $\overline{\Sigma}$ of Σ in $(\mathbb{P}^1)^n$ satisfies $\overline{\Sigma} \neq (\mathbb{P}^1)^n$ and the Zariski closure $\overline{\Gamma^{\mathsf{T}}}$ of Γ^{T} in GL_2^n satisfies $\overline{\Gamma^{\mathsf{T}}}\, z \subset \overline{\Sigma}$. However, $\overline{\Gamma^{\mathsf{T}}} \supset \mathrm{SL}_2^n$ by (10). We have a contradiction since $\mathrm{SL}_2(\mathbb{C})^n$ acts transitively on $(\mathbb{P}^1)^n(\mathbb{C})$. □

Proof of Proposition 3.1. Let C and $z \in j^{-1}(C(\mathbb{C}))$ be as in the hypothesis. The holomorphic map $j : \mathbb{H}^n \to \mathbb{C}^n$ is open since each factor is open. So the restriction $j|_{j^{-1}(C(\mathbb{C}))} : j^{-1}(C(\mathbb{C})) \to C(\mathbb{C})$ is open. It is also finite at all points of its domain by the comment on page 64 of [GR84] and because j has discrete fibers. By the corollary on page 105 of [GR84] we find $\dim_z j^{-1}(C(\mathbb{C})) = 1$.

The map given by (1) determines a morphism of affine curves $Y(2) \to Y(1)$. This morphism is quasi-finite, i.e. its fibers are finite. Therefore, the n-fold product $f : Y(2)^n \to Y(1)^n$ is also quasi-finite. Say $C_1 \cup \cdots \cup C_r$ is the decomposition of $f^{-1}(C)$ into irreducible components. By Lemma 3.1(iii) we have $j^{-1}(C(\mathbb{C})) = u^{-1}(C_1(\mathbb{C})) \cup \cdots \cup u^{-1}(C_r(\mathbb{C}))$. Without loss of generality we may assume $\dim_z u^{-1}(C_1(\mathbb{C})) = 1$. After replacing z by a sufficiently close point we may also assume that z is a smooth point of $u^{-1}(C_1(\mathbb{C}))$ and that some neighborhood of z in $u^{-1}(C_1(\mathbb{C}))$ is in a hypersurface of $\mathbb{H}^n$. Clearly, we must have $\dim C_1 \geqslant 1$. However, f is quasi-finite and $f(C_1) \subset C$, so C_1 is a curve too. The result follows from Lemma 3.4 applied to C_1. □

4. Lower bounds for Galois orbits

In this section we establish lower bounds for Galois orbits of points in the intersection of a curve with $\mathcal{S}^{[2]}$ and $\mathcal{S}^{\mathrm{nss},[2]}$.

Our first lemma relies on the André–Oort conjecture for curves in $Y(1)^2$. This result is known unconditionally by the work of André [And98] or by the second author [Pil09b].

Lemma 4.1. *Let $C \subset Y(1)^n$ be an irreducible curve defined over $\overline{\mathbb{Q}}$ that is not contained in a special subvariety of positive codimension. Then*

$$\{(x_1, \ldots, x_n) \in C(\overline{\mathbb{Q}});\ \text{there are } 1 \leqslant i \neq i' \leqslant n \text{ with } x_i \text{ and } x_{i'} \text{ special}\} \tag{12}$$

is finite.

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

Proof. We may assume $n \geqslant 2$. For a fixed pair $1 \leqslant i < i' \leqslant n$ we let $C' \subset Y(1)^2$ denote the Zariski closure of the projection of C onto the ith and i'th coordinate. Then C' is irreducible and $\dim C' \leqslant 1$. If C' contains infinitely many special points, then $\dim C' = 1$ and by the André–Oort conjecture it is either a horizontal line, a vertical line, or the set of zeros of Φ_N for some N. In any of these cases, C is contained in a special subvariety of positive codimension. □

The next lemma gives a lower bound for Galois orbits for points in $C \cap \mathcal{S}^{[2]}$ if C is an asymmetric curve. If one could drop the assumption on asymmetry here, even at the cost of replacing $1/6$ by a smaller positive constant, then one would be able to drop the assumption on asymmetry in Theorem 1.

Lemma 4.2. *Let $C \subset Y(1)^n$ be an irreducible asymmetric curve defined over $\overline{\mathbb{Q}}$ that is not contained in a special subvariety of positive codimension. Let Σ be the finite set (12). There exists a constant $c = c(C) > 0$ with the following property. Let $i_1, i_2, i_3, i_4 \in \{1, \dots, n\}$ with $i_1 \neq i_2$ and $i_3 \neq i_4$; if $X_{i_1}|_C$ and $X_{i_2}|_C$ are both non-constant we shall additionally assume $\{i_1, i_2\} \neq \{i_3, i_4\}$. If $x = (x_1, \dots, x_n) \in C(\overline{\mathbb{Q}}) \smallsetminus \Sigma$ such that there exist positive integers M and N with $\Phi_M(x_{i_1}, x_{i_2}) = \Phi_N(x_{i_3}, x_{i_4}) = 0$, then*

$$[\mathbb{Q}(x) : \mathbb{Q}] \geqslant c \max\{M, N\}^{1/6}.$$

Proof. Below, $c_1, c_2, \dots$ denote positive constants which may depend on C but not on x, M, and N. They will determine our final constant c. Let $i_1, \dots, i_4, x = (x_1, \dots, x_n), M$, and N be as in the hypothesis. We write $d_k = \deg(X_{i_k}|_C)$ for $k = 1, \dots, 4$.

We remark that $d_1 = d_2 = 0$ is impossible. Indeed, otherwise $X_{i_1}|_C$ and $X_{i_2}|_C$ would be constant with values x_{i_1} and x_{i_2}, respectively. Now $\Phi_M(x_{i_1}, x_{i_2}) = 0$ implies that C is contained in a special subvariety of positive codimension, which is a contradiction. For the same reason we cannot have $d_3 = d_4 = 0$. If d_1 or d_2 is zero, then $d_1 \neq d_2$. If both are non-zero then $\{i_1, i_2\} \neq \{i_3, i_4\}$ by hypothesis. After possibly permuting $\{i_1, i_2\}$ and $\{i_3, i_4\}$ we may assume $d_1 \neq d_2$ since our curve is asymmetric. Hence we have $d_1 \neq d_2$ in any case; we may even assume $d_1 \geqslant d_2 + 1 \geqslant 1$ after exchanging i_1 and i_2.

We define $C' \subset Y(1)^2$ as the Zariski closure of the projection of C onto the i_1st and i_2nd coordinates. Then C' is an irreducible variety of dimension at most one. However, it must be a curve since $X_{i_1}|_C$ is non-constant. We use X and Y to denote the two coordinate functions on $Y(1)^2$. By considering function fields, the pair (d_1, d_2) is $(\deg(X|_{C'}), \deg(Y|_{C'}))d_0$ with d_0 a positive integer. Hence C' is cut out by an irreducible polynomial $A \in \overline{\mathbb{Q}}[X, Y]$ with $\deg_Y A = \deg(X|_{C'}) = d_1/d_0$ and $\deg_X A = \deg(Y|_{C'}) = d_2/d_0$.

By hypothesis we have $\Phi_M(x_{i_1}, x_{i_2}) = 0$. Before coming to degree lower bounds we need a sufficiently strong height upper bound for $h(x_{i_1})$ and $h(x_{i_2})$ in terms of M. To do this we follow the lines of [Hab10, Proof of Theorem 1.1].

It is known from the work of Siegel and Néron that

$$d_0^{-1}|h(x_{i_1})d_2 - h(x_{i_2})d_1| = |h(x_{i_1})\deg_X A - h(x_{i_2})\deg_Y A| \leqslant c_1 \max\{1, h(x_{i_1}), h(x_{i_2})\}^{1/2}; \quad (13)$$

for a proof see [Hab07, Theorem A.1]. (The case $\deg_X A = 0$ is formally not covered by this reference. If $\deg_X A = 0$, then we may take $A = Y - x_{i_2}$ and so $d_2 = 0$. In this case inequality (13) is easy to show.)

On the other hand, there are elliptic curves E_1 and E_2 defined over the number field $\mathbb{Q}(x_{i_1}, x_{i_2})$ with j-invariants x_{i_1} and x_{i_2}, respectively. Let $h_F(E_1)$ and $h_F(E_2)$ denote their semi-stable Faltings height. Silverman's comparison estimate between the Faltings height and the height of

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

the j-invariant implies

$$|h(x_{i_1}) - 12h_F(E_1)| \leqslant c_2 \log \max\{2, h(x_{i_1})\}. \tag{14}$$

The same inequality holds for the pair x_{i_2}, E_2. Moreover, $\Phi_M(x_{i_1}, x_{i_2}) = 0$ is equivalent to saying that there is an isogeny with cyclic kernel of order M between E_1 and E_2. By a result of Faltings, cf. [Ray85, Corollary 2.1.4], we have

$$|h_F(E_1) - h_F(E_2)| \leqslant \tfrac{1}{2} \log M.$$

This estimate and (14) imply

$$\begin{aligned} |h(x_{i_1}) - h(x_{i_2})| &= |h(x_{i_1}) - 12h_F(E_1) - (h(x_{i_2}) - 12h_F(E_2)) + 12(h_F(E_1) - h_F(E_2))| \\ &\leqslant |h(x_{i_1}) - 12h_F(E_1)| + |h(x_{i_2}) - 12h_F(E_2)| + 12|h_F(E_1) - h_F(E_2)| \\ &\leqslant 2c_2 \log \max\{2, h(x_{i_1}), h(x_{i_2})\} + 6 \log M. \end{aligned} \tag{15}$$

The inequalities $d_1 \geqslant d_2 + 1$ and $h(x_{i_1}) \geqslant 0$ give

$$\begin{aligned} h(x_{i_1}) &\leqslant h(x_{i_1})(d_1 - d_2) \\ &= h(x_{i_1})d_1 - h(x_{i_2})d_1 + h(x_{i_2})d_1 - h(x_{i_1})d_2 \\ &\leqslant |h(x_{i_1}) - h(x_{i_2})|d_1 + |h(x_{i_2})d_1 - h(x_{i_1})d_2|. \end{aligned}$$

We insert (15) and (13) into this inequality to obtain

$$h(x_{i_1}) \leqslant c_3(\log M + \max\{1, h(x_{i_1}), h(x_{i_2})\}^{1/2}). \tag{16}$$

Since $d_1 \geqslant 1$ we may use (13) again to deduce $h(x_{i_2}) \leqslant c_4(1 + h(x_{i_1}))$. Hence the contribution of $h(x_{i_2})$ to the bound in (16) is harmless. We obtain

$$h(x_{i_1}) \leqslant c_5 \max\{1, \log M\}. \tag{17}$$

We apply (14) again to conclude

$$h_F(E_1) \leqslant c_6 \max\{1, \log M\}. \tag{18}$$

Enter isogeny estimates. By Pellarin's Théorème 2 [Pel01] there exists an isogeny $E_1 \to E_2$ of degree M' with

$$M' \leqslant c_7[\mathbb{Q}(x_{i_1}, x_{i_2}) : \mathbb{Q}]^5 \max\{1, h_F(E_1)\}^2.$$

Using the bound (18) leads to

$$M' \leqslant c_8[\mathbb{Q}(x_{i_1}, x_{i_2}) : \mathbb{Q}]^5 \max\{1, \log M\}^2. \tag{19}$$

We remark that any upper bound for M' which is polynomial in the product

$$[\mathbb{Q}(x_{i_1}, x_{i_2}) : \mathbb{Q}] \max\{1, h_F(E_1)\}$$

leads to a version of this lemma which is strong enough to imply Theorem 1. For example, instead of Pellarin's bound we could also apply a variant of David's earlier estimate [Dav95].

Recall that $x \notin \Sigma$, so E_1 and E_2 cannot both have complex multiplication. However, these elliptic curves are isogenous, so neither has complex multiplication. The group of all homomorphisms $E_1 \to E_2$ is generated by any cyclic isogeny, cf. [Hab10, Lemma 3.2]. In particular, $M \leqslant M'$. Inequality (19) implies $M \leqslant c_8[\mathbb{Q}(x_{i_1}, x_{i_2}) : \mathbb{Q}]^5 \max\{1, \log M\}^2$. A simple calculation leads to

$$M \leqslant c_9[\mathbb{Q}(x_{i_1}, x_{i_2}) : \mathbb{Q}]^6 \leqslant c_9[\mathbb{Q}(x) : \mathbb{Q}]^6 \tag{20}$$

and this is the first half of our claim. It remains to find a similar bound for N.

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

Let $C'' \subset Y(1)^2$ be the Zariski closure of the projection of C to the i_1st and i_3rd coordinates. Since $d_1 \neq 0$ we see that C'' is an irreducible curve. Using an argument similar to the one given around (13) we deduce $|d_1 h(x_{i_3}) - d_3 h(x_{i_1})| \leqslant c_{10} \max\{1, h(x_{i_1}), h(x_{i_3})\}^{1/2}$. Hence

$$h(x_{i_3}) \leqslant c_{11} \max\{1, h(x_{i_1})\} \leqslant c_{12} \max\{1, \log M\} \tag{21}$$

follows from (17) and $d_1 \geqslant 1$. As above, there are elliptic curves E_3 and E_4 defined over $\mathbb{Q}(x_{i_3}, x_{i_4})$ whose j-invariants equal x_{i_3} and x_{i_4}, respectively. Just as above, we deduce that E_3 and E_4 cannot have complex multiplication. Since $\Phi_N(x_{i_3}, x_{i_4}) = 0$ there is an isogeny $E_3 \to E_4$ with cyclic kernel of order N. Using Silverman's result, comparing $h(x_{i_3})$ and $12h_F(E_3)$ together with (21) leads to $h_F(E_3) \leqslant c_{13} \max\{1, \log M\}$. We apply Pellarin's isogeny estimate in a similar manner as above to deduce $N \leqslant c_{14}[\mathbb{Q}(x_{i_3}, x_{i_4}) : \mathbb{Q}]^5 \max\{1, \log M\}^2 \leqslant c_{14}[\mathbb{Q}(x) : \mathbb{Q}]^5 \max\{1, \log M\}^2$. This is nearly the equality we are looking for; it remains to treat the factor $\max\{1, \log M\}^2$. However, this is at most $c_{15}[\mathbb{Q}(x) : \mathbb{Q}]$ by (20). □

If $x \in Y(1)(\mathbb{C})$ we define $\Delta(x) \in \mathbb{Z}$ to be the discriminant of the endomorphism ring of an elliptic curve with j-invariant equal to x. For example, if x is special then $\Delta(x) \leqslant -3$ and if x is not special then $\Delta(x) = 1$.

In Lemma 4.4 below we give a lower bound for Galois orbits of points in the intersection of a curve with $\mathcal{S}^{\mathrm{nss},[2]}$, the union of all non-strongly special subvarieties of codimension at least two. The presence of one coordinate which is special enables us to drop the condition on asymmetry. Indeed, the following lemma controls the height of this coordinate. Its proof relies on a generalization of the Chowla–Selberg formula by Nakkajima and Taguchi [NT91] and a classical estimate for Siegel zeros. The referee has pointed out that the following estimate follows from a result of Breuer [Bre01] if $\epsilon = 1/2$.

Lemma 4.3. *Let $\epsilon > 0$, there is a constant $c = c(\epsilon) > 0$ such that if $x \in Y(1)(\overline{\mathbb{Q}})$ is special then $h(x) \leqslant c|\Delta(x)|^{\epsilon}$.*

Proof. In this proof $c_1, c_2, \ldots$ denote positive constants that depend only on ϵ.

Let E be an elliptic curve with j-invariant x. We may write $\Delta(x) = f^2\Delta$ where $\Delta < 0$ is the fundamental discriminant and $f \geqslant 1$ an integer. Let $\chi(\cdot) = (\frac{\Delta}{\cdot})$ be the Kronecker symbol and $L(\chi, s)$ the Dirichlet L-function associated to χ. The latter is an entire function and non-zero at $s = 1$.

A consequence of Nakkajima and Taguchi's aforementioned result is

$$h_F(E) = \frac{1}{4}\log(f^2|\Delta|) + \frac{1}{2}\frac{L'(\chi,1)}{L(\chi,1)} - \frac{1}{2}\sum_{p|f} e(p)\log p - \frac{\gamma}{2} - \frac{\log(2\pi)}{2}$$

where $h_F(E)$ is again the semi-stable Faltings height of E, γ is Euler's constant, and $e(p) \geqslant 0$; cf. [Hab10, Lemma 4.1]. Hence

$$h_F(E) \leqslant \frac{1}{4}\log|\Delta(x)| + \frac{1}{2}\frac{L'(\chi,1)}{L(\chi,1)}. \tag{22}$$

By [GS00, Remark 1, p. 515] we have

$$\frac{L'(\chi,1)}{L(\chi,1)} \leqslant \frac{1}{1-\beta} + c_1 \log|\Delta| \tag{23}$$

where $\beta \in (0,1)$ is a Siegel zero, if such a thing exists; if not then (23) holds without the term $(1-\beta)^{-1}$. In the former case, [Dav00, Siegel's theorem, p. 126], tells us that there is c_2 such

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

that $\beta \leqslant 1 - c_2|\Delta|^{-\epsilon}$. This bound together with (22) and (23) enables us to deduce $h_F(E) \leqslant c_3|\Delta_E|^{\epsilon}$ regardless of whether a Siegel zero exists or not.

Silverman's bound for $|h(x) - 12h_F(E)|$, which already appear above in (14), completes the proof. □

Lemma 4.4. *Let $C \subset Y(1)^n$ be an irreducible curve defined over $\overline{\mathbb{Q}}$ that is not contained in a special subvariety of positive codimension. Let Σ be the finite set (12). There exists a constant $c = c(C) > 0$ with the following property. Let $i_1, i_2, i_3 \in \{1, \dots, n\}$ be pairwise distinct. If $x = (x_1, \dots, x_n) \in C(\overline{\mathbb{Q}}) \smallsetminus \Sigma$ such that x_{i_1} is special and such that there is a positive integer N with $\Phi_N(x_{i_2}, x_{i_3}) = 0$, then*

$$[\mathbb{Q}(x) : \mathbb{Q}] \geqslant c \max\{|\Delta(x_{i_1})|, N\}^{1/6}.$$

Proof. In this proof $c_1, c_2, \dots$ denote positive constants that depend only on C but not on x. Let $i_1, i_2, i_3, x = (x_1, \dots, x_n)$, and N be as in the hypothesis.

There is an irreducible polynomial $A \in \overline{\mathbb{Q}}[X, Y]$ such that $A(X_{i_1}|_C, X_{i_2}|_C) = 0$. We have $\deg_Y A \geqslant 1$. Indeed, otherwise A would be linear and in one variable X. Then the i_1st coordinate of any point in $C(\overline{\mathbb{Q}})$ would be x_{i_1}. This is impossible since x_{i_1} is special and because C is not contained in a special subvariety of positive codimension.

As in (13) we have $|h(x_{i_1}) \deg_X A - h(x_{i_2}) \deg_Y A| \leqslant c_1 \max\{1, h(x_{i_1}), h(x_{i_2})\}^{1/2}$. Since $\deg_Y A \geqslant 1$ we may estimate $h(x_{i_2}) \leqslant c_2 \max\{1, h(x_{i_1})\}$. However, x_{i_1} is special by hypothesis and so Lemma 4.3 with $\epsilon = 1/6$ yields

$$h(x_{i_2}) \leqslant c_3|\Delta(x_{i_1})|^{1/6}. \tag{24}$$

The algebraic numbers x_{i_2} and x_{i_3} are related by $\Phi_N(x_{i_2}, x_{i_3}) = 0$. Moreover, neither one is special since we are assuming $(x_1, \dots, x_n) \notin \Sigma$.

We now follow a similar line of argumentation as around (19) and thus give less detail. First we apply Silverman's height comparison result (14). We obtain from (24) an upper bound, linear in $|\Delta(x_{i_1})|^{1/6}$, for the semi-stable Faltings height of an elliptic curve with j-invariant x_{i_2}. We then apply Pellarin's isogeny estimate to bound the degree of an isogeny between elliptic curves with j-invariant equal to x_{i_2} and x_{i_3}. Because these elliptic curves do not have complex multiplication we obtain

$$N \leqslant c_4[\mathbb{Q}(x_{i_2}, x_{i_3}) : \mathbb{Q}]^5|\Delta(x_{i_1})|^{1/3}. \tag{25}$$

The degree $[\mathbb{Q}(x_{i_1}) : \mathbb{Q}]$ is the class number of the endomorphism ring of an elliptic curve with j-invariant x_{i_1}, cf. [Cox89, ch.s 10 and 11]. The discriminant of this ring is $\Delta(x_{i_1}) = f^2\Delta$ with $\Delta < 0$ a fundamental discriminant and $f \geqslant 1$ an integer. Using [Cox89, Theorem 7.24] we can express

$$[\mathbb{Q}(x_{i_1}) : \mathbb{Q}] = \frac{Gf}{w} \prod_{p|f} \left(1 - \left(\frac{\Delta}{p}\right)\frac{1}{p}\right)$$

where G is the class number of $\mathbb{Q}(\sqrt{\Delta})$, $w \in \{1, 2, 3\}$, and p runs over the prime divisors of f. Now

$$f \prod_{p|f} \left(1 - \left(\frac{\Delta}{p}\right)\frac{1}{p}\right) \geqslant f \prod_{p|f} \left(1 - \frac{1}{p}\right) = \varphi(f),$$

where φ is Euler's totient function. It is well-known that $\varphi(f) \geqslant c_5 f^{2/3}$. Using Siegel's theorem, cf. [Dav00, ch. 21], we can bound $G \geqslant c_6|\Delta|^{(1/2)-(1/6)}$; this crude bound will be perfectly sufficient for our purposes.

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

Putting these estimates together leads to $|\Delta(x_{i_1})| \leqslant c_7[\mathbb{Q}(x_{i_1}):\mathbb{Q}]^3$. First we use it to obtain the upper bound for $|\Delta(x_{i_1})|$ in the assertion. However, we can also combine it with (25) to get

$$N \leqslant c_4 c_7^{1/3}[\mathbb{Q}(x_{i_2}, x_{i_3}):\mathbb{Q}]^5[\mathbb{Q}(x_{i_1}):\mathbb{Q}] \leqslant c_8[\mathbb{Q}(x):\mathbb{Q}]^6. \qquad \square$$

5. Synthesis

In this section, if not stated otherwise, definable will mean definable in the o-minimal structure $\mathbb{R}_{\mathrm{an,exp}}$.

5.1 Technicalities

Recall that the height of $\alpha \in \mathrm{Mat}_2(\mathbb{Q})$ is its height when regarded as an element of $\mathbb{Q}^4$. If $\alpha, \beta \in \mathrm{Mat}_2(\mathbb{Z})$, then the triangle inequality implies $H(\alpha\beta) \leqslant 2H(\alpha)H(\beta)$. We often use this elementary inequality without further mention.

If $z \in \mathbb{H}$ is algebraic, then, since we have identified $\mathbb{H}$ with a subset of $\mathbb{R}^2$ the height $H(z)$ is by definition $\max\{H(\mathrm{Re}(z)), H(\mathrm{Im}(z))\}$.

The next lemma will be used to effectively bring any element of $\mathbb{H}$ into the fundamental domain F. Its elementary proof uses the function

$$D(z) = \max\{1, |\mathrm{Re}(z)|, \mathrm{Im}(z)^{-1}\} \quad \text{for } z \in \mathbb{H}.$$

The result is slightly stronger than [Pil11, Proposition 5.2], which gives a similar bound in terms of $D^*(z) = \max\{1, |z|, \mathrm{Im}(z)^{-1}\}$, which would suffice for the immediate purposes.

Lemma 5.1. *There exists an absolute constant $c > 0$ with the following property. If $z \in \mathbb{H}$ there is $\rho \in \mathrm{SL}_2(\mathbb{Z})$ with $H(\rho) \leqslant cD(z)^9$ such that $\rho z \in F$.*

Proof. Say $z = x + yi$ with $x, y \in \mathbb{R}$ and $y > 0$. Let us first assume $y \leqslant 1/2$. Let us define $Q = y^{-1/2} > 1$. By a classical result of Dirichlet, see [Cas57, p. 1], there are integers $c, d \in \mathbb{Z}$ with $1 \leqslant c < Q$ and $|cx + d| \leqslant Q^{-1}$. Without loss of generality we may assume that c and d are coprime. There are integers $a, b \in \mathbb{Z}$ with $\max\{|a|, |b|\} \leqslant \max\{|c|, |d|\}$ and $ad - bc = 1$. We define $\beta = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ and note $H(\beta) = \max\{|c|, |d|\}$. Now $|c| < y^{-1/2} \leqslant D(z)^{1/2}$ and $|d| \leqslant |c||x| + Q^{-1} \leqslant D(z)^{1/2}|x| + 1 \leqslant 2D(z)^{3/2}$, so

$$H(\beta) \leqslant 2D(z)^{3/2}. \tag{26}$$

We have $|\beta z| = |az + b|/|cz + d|$. If $c \neq 0$, then $|cz + d| \geqslant |\mathrm{Im}(cz + d)| = |c|\mathrm{Im}(z) \geqslant D(z)^{-1}$. However, $|cz + d| \geqslant D(z)^{-1}$ also holds for $c = 0$. We have the useful inequality $|z|^2 \leqslant x^2 + 1/4 \leqslant 2D(z)^2$. Now, $|az + b| \leqslant |a||z| + |b| \leqslant 2|a|D(z) + |b| \leqslant 3H(\beta)D(z)$ and so combining the bounds from above we get $|\beta z| \leqslant 3H(\beta)D(z)^2 \leqslant 6D(z)^{7/2}$. For the imaginary part we have $\mathrm{Im}(\beta z) = \mathrm{Im}(z)|cz + d|^{-2} = y((cx + d)^2 + c^2y^2)^{-1} \geqslant y(Q^{-2} + Q^2y^2)^{-1} = y(y + y)^{-1}$, so

$$\mathrm{Im}(\beta z) \geqslant \tfrac{1}{2}. \tag{27}$$

We use $|cz + d| \geqslant D(z)^{-1}$ again to deduce

$$|\mathrm{Re}(\beta z)| = \frac{|bd + x(ad + bc) + ac|z|^2|}{|cz + d|^2} \leqslant 4H(\beta)^2D(z)^2 \max\{1, |x|, |z|^2\}. \tag{28}$$

This leads to $|\mathrm{Re}(\beta z)| \leqslant 8H(\beta)^2D(z)^4$. We recall (26) to conclude

$$|\mathrm{Re}(\beta z)| \leqslant 32D(z)^7. \tag{29}$$

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

Now if $y > 1/2$, then (26), (27), and (29) hold with β the identity matrix.

There is one $b' \in \mathbb{Z}$ with $\mathrm{Re}(\beta z) + b' \in [-1/2, 1/2)$. Let $\gamma = \left[\begin{smallmatrix} 1 & b' \\ 0 & 1 \end{smallmatrix}\right]$, then $\mathrm{Re}(\gamma\beta z) \in [-1/2, 1/2)$ and $\mathrm{Im}(\gamma\beta z) \geqslant 1/2$. Now $|b'| \leqslant |\mathrm{Re}(\beta z)| + 1/2 \leqslant 33D(z)^7$ by (29), so $H(\gamma) = \max\{1, |b'|\} \leqslant 33D(z)^7$.

Finally, we conclude that $\delta\gamma\beta z \in F$ for δ one of

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad \text{or} \quad \begin{bmatrix} \pm 1 & -1 \\ 1 & 0 \end{bmatrix};$$

of course, $H(\delta) = 1$.

It follows that the product $\rho = \delta\gamma\beta \in \mathrm{SL}_2(\mathbb{Z})$ satisfies $\rho z \in F$. We have $H(\rho) \leqslant 2H(\delta)H(\gamma\beta) \leqslant 4H(\gamma)H(\beta) \leqslant 264D(z)^{17/2}$, using the height estimates derived above. □

The precise bound for $H(\rho)$ above is not so important. For our applications it suffices to bound the height polynomially in $D(z)$.

Lemma 5.2. *There exists an absolute constant $c > 0$ with the following property. Let $x_1, x_2 \in Y(1)(\mathbb{C})$ with $\Phi_N(x_1, x_2) = 0$ for some integer $N \geqslant 1$ and $z_1, z_2 \in F$ with $j(z_i) = x_i$. There exist $\alpha \in \mathrm{Mat}_2(\mathbb{Z})$ such that*

$$\det \alpha = N, \quad z_2 = \alpha z_1 \quad \text{and} \quad H(\alpha) \leqslant cN^{10}.$$

Proof. By [Lan87, ch. 5.2] there are coprime and non-negative integers a, b, d with $ad = N$ and $0 \leqslant b < d$ such that z_2 and $z = (az_1 + b)/d$ are in the same $\mathrm{SL}_2(\mathbb{Z})$-orbit. Since $z_1 \in F$ we have $|\mathrm{Re}(z_1)| \leqslant 1/2$ and $\mathrm{Im}(z_1) \geqslant 1/2$. We bound $|\mathrm{Re}(z)| \leqslant a|\mathrm{Re}(z_1)|/d + b/d \leqslant a/(2d) + 1 \leqslant N/2 + 1 \leqslant 2N$. Next we estimate the imaginary part $\mathrm{Im}(z) = a\,\mathrm{Im}(z_1)/d \geqslant a/(2d) \geqslant 1/(2N)$. Hence $D(z) \leqslant 2N$. By Lemma 5.1 there is an absolute constant $c' > 0$ and $\rho \in \mathrm{SL}_2(\mathbb{Z})$ with $\rho z \in F$ and $H(\rho) \leqslant c'D(z)^9 \leqslant 2^9 c' N^9$. However, the elements ρz and z_2 of the fundamental domain are in the same $\mathrm{SL}_2(\mathbb{Z})$-orbit, so $z_2 = \rho z$. This lemma follows with

$$\alpha = \rho \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$$

since $H(\alpha) \leqslant 2H(\rho)\max\{a, b, d\} \leqslant c'2^{10}N^{10}$. □

We recall that if $x \in Y(1)(\mathbb{C})$ is special, then $\Delta(x)$ is the discriminant of the endomorphism ring of an elliptic curve with j-invariant equal to x. The estimate in the following lemma is a variant of one in [Pil11, Proposition 5.7].

Lemma 5.3. *Let $z \in F$ be special. Then $[\mathbb{Q}(z) : \mathbb{Q}] = 2$ and $H(z) \leqslant 2|\Delta(j(z))|$.*

Proof. The statement on the degree of z over $\mathbb{Q}$ follows by definition.

Let $a, b, c \in \mathbb{Z}$ be coprime integers with $a \geqslant 1$ and $az^2 + bz + c = 0$. It follows from [Lan87, Theorem 1, p. 90] that $|\Delta(j(z))| = 4ac - b^2 > 0$. Now $\mathrm{Re}(z) = -b/(2a)$ and $\mathrm{Im}(z) = |\Delta(j(z))|^{1/2}/(2a)$. The fact that $z \in F$ implies $|\mathrm{Re}(z)| \leqslant 1/2$ and $\mathrm{Im}(z) \geqslant 1/2$ and hence

$$|b| \leqslant a \leqslant |\Delta(j(z))|^{1/2}. \tag{30}$$

By definition of the height we have $H(\mathrm{Re}(z)) \leqslant \max\{2a, |b|\}$ and $H(\mathrm{Im}(z)) \leqslant 2a|\Delta(j(z))|^{1/2}$. We now apply (30) to deduce $H(\mathrm{Re}(z)) \leqslant 2|\Delta(j(z))|^{1/2}$ and $H(\mathrm{Im}(z)) \leqslant 2|\Delta(j(z))|$; this implies the desired bound for $H(z)$. □

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

5.2 Two pairs of linked coordinates

Proposition 5.1. *Let $C \subset Y(1)^n$ be an irreducible asymmetric curve defined over $\overline{\mathbb{Q}}$ that is not contained in a special subvariety of positive codimension. Let $i_1, i_2, i_3, i_4 \in \{1, \dots, n\}$ with $i_1 \neq i_2$, $i_3 \neq i_4$ and $\{i_1, i_2\} \neq \{i_3, i_4\}$. There are only finitely many $(x_1, \dots, x_n) \in C(\overline{\mathbb{Q}})$ for which there exist positive integers M and N with*

$$\Phi_M(x_{i_1}, x_{i_2}) = \Phi_N(x_{i_3}, x_{i_4}) = 0. \tag{31}$$

Proof. Let

$$Z = j^{-1}(C(\mathbb{C})) \cap F^n.$$

Then Z is definable. For each $\alpha, \beta \in \mathrm{GL}_2(\mathbb{R})^+$ the set

$$Y_{\alpha,\beta} = \{(z_1, \dots, z_n) \in \mathbb{H}^n;\, z_{i_2} = \alpha z_{i_1} \text{ and } z_{i_4} = \beta z_{i_3}\}$$

is definable; indeed it is semi-algebraic in the real coordinates.

We claim we may assume that $Y_{\alpha,\beta} \cap Z$ is finite for all possible α, β. The difficulty here lies in the fact that α and β need not have rational coefficients. Indeed, assuming the contrary, there are α and β with $Y_{\alpha,\beta} \cap Z$ infinite. This intersection is definable and has positive dimension. Hence it contains uncountably many points. The larger set $Y_{\alpha,\beta} \cap j^{-1}(C(\mathbb{C}))$ is certainly uncountable as well. However, this intersection is now a complex analytic subset of $\mathbb{H}^n$; as such, it contains a smooth z point of dimension at least one. Using the same argument as in the proof of Proposition 3.1 we find that the complex analytic space $j^{-1}(C(\mathbb{C}))$ has dimension one at all points. By comparing dimensions, some neighborhood of z in $j^{-1}(C(\mathbb{C}))$ is contained in $Y_{\alpha,\beta} \cap j^{-1}(C(\mathbb{C}))$. The algebraic relation $z_{i_2} = \alpha z_{i_1}$ holds on this neighborhood. Let $C' \subset Y(1)^2$ be the projection of C onto the i_1st and i_2nd coordinates of $Y(1)^n$. Surely, $\dim C' \leqslant 1$. If $\dim C' = 0$ then C cannot contain any point satisfying the first equality (31) simply because it is not contained in a special subvariety of positive codimension. The proposition holds in this case. Hence we may assume $\dim C' = 1$. By the argument above, $j^{-1}(C'(\mathbb{C}))$ contains a neighborhood inside a hypersurface of $\mathbb{H}^n$. We invoke Proposition 3.1 to conclude that C' is inside a geodesic subvariety of positive codimension. This geodesic subvariety cannot be special by hypothesis. Therefore, either the i_1st or the i_2nd coordinate is constant on C'. The assumption that $Y_{\alpha,\beta} \cap Z$ is infinite forces the other coordinate to be constant too. Hence $\dim C' = 0$, which is a contradiction.

Now $Y_{\alpha,\beta}$ is a fiber of the definable family

$$Y = \{(z_1, \dots, z_n, \alpha, \beta) \in \mathbb{H}^n \times \mathrm{GL}_2(\mathbb{R})^+ \times \mathrm{GL}_2(\mathbb{R})^+;\, z_{i_2} = \alpha z_{i_1}, z_{i_4} = \beta z_{i_3}\}.$$

By o-minimality there is a uniform finite bound for the number of points in any $Y_{\alpha,\beta} \cap Z$. Consider the definable set

$$X = \{(\alpha, \beta) \in \mathrm{GL}_2(\mathbb{R})^+ \times \mathrm{GL}_2(\mathbb{R})^+;\, Y_{\alpha,\beta} \cap Z \neq \emptyset\}.$$

Suppose now that $x = (x_1, \dots, x_n) \in C(\overline{\mathbb{Q}})$ as in the hypothesis. Excluding a finite number of points as in Lemma 4.1, we may suppose that $x_{i_1}, x_{i_2}, x_{i_3}, x_{i_4}$ are not special. We use

$$L = \max\{M, N\}$$

to measure the complexity of the special subvariety containing x. In the following, $\delta = 1/6$, and $c_1, c_2, \dots$ denote positive constants that may depend on C but not on M, N, and are not related to earlier choices of these constants.

In order to prove the proposition, it is enough to show that L is bounded in terms of C only. Hence we will assume that L is sufficiently large and aim at a contradiction.

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

According to Lemma 4.2, x has at least

$$c_1 L^{\delta}$$

conjugates x' over $\mathbb{Q}$, and some fixed proportion, depending only on the degree over $\mathbb{Q}$ of a field of definition for C, of these lie again on C. The conjugates $x' = (x'_1, \dots, x'_n)$ are also points with $\Phi_M(x'_{i_1}, x'_{i_2}) = \Phi_N(x'_{i_3}, x'_{i_4}) = 0$ and so lead to the same complexity.

Each such x' has a pre-image $(z_1, \dots, z_n) \in Z$, and by Lemma 5.2 we can find $\alpha, \beta \in \mathrm{GL}_2(\mathbb{Q})^+$ with height at most

$$c_2 L^{c_3}$$

such that $z_{i_2} = \alpha z_{i_1}$ and $z_{i_4} = \beta z_{i_3}$. The rational point (α, β) is in X.

Therefore, X contains at least $c_4 L^{\delta}$ rational points up to height $c_2 L^{c_3}$, corresponding to at least $c_4 L^{\delta}$ distinct intersections of the corresponding geodesic subvarieties with Z. By Theorem 5 with $\epsilon = \delta/(2c_3)$, the rational points on X up to height $c_2 L^{c_3}$ lie on at most

$$c_5 L^{\delta/2}$$

blocks $W_z^{(j)}$ from a finite number of block families $W^{(j)}$ with $j = 1, \dots, J$. These block families depend only on X and the choice of $\epsilon = \delta/(2c_3)$.

Let $W_z^{(j)}$ be a fiber of one of these families. Each point $(\alpha, \beta) \in W_z^{(j)}$ corresponds to a set $Y_{\alpha,\beta}$ whose intersection with Z is a set of uniformly bounded cardinality, independently of α, β.

For $j = 1, \dots J$ we can find finitely many definable functions $f_{z,i}^{(j)} : W_z^{(j)} \to Y_{\alpha,\beta} \cap Z$; here i runs over a finite index set depending on j, definable as a family of functions over all the blocks in the family $W^{(j)}$, that parameterize all the points in the intersections $Y_{\alpha,\beta} \cap Z$. That is, for every j, every z in the parameter set for $W^{(j)}$ and all $(\alpha, \beta) \in W_z^j$ we have $\bigcup_i f_{z,i}^{(j)}(\alpha, \beta) = Y_{\alpha,\beta} \cap Z$. Every $Y_{\alpha,\beta}$ with $(\alpha, \beta) \in W_z^{(j)}$ intersects Z, so we may assume the $f_{z,i}^{(j)}$ are defined on all $W_z^{(j)}$. They are then once continuously differentiable outside some lower dimensional subset $S_z^{(j)}$, which is the fiber of a definable family $S^{(j)}$. If $W_z^{(j)} \smallsetminus S_z^{(j)}$ has any points that are not regular of highest dimension they constitute a set of smaller dimension than $W_z^{(j)}$, definable over the family, and we add them to $S_z^{(j)}$.

Let now $W_z^{(j)}$ be one of the blocks afforded by Theorem 5 containing rational points of X. Suppose there is a point in $W_z^{(j)} \smallsetminus S_z^{(j)}$ where the differential of one of the $f_{z,i}^{(j)}$ is non-zero. Then there is a semi-algebraic curve in $W_z^{(j)}$ on which the intersection with Z is non-constant.

We soon split up into cases depending on the set I of indices $\{i_1, i_2, i_3, i_4\}$ for which the corresponding coordinate function is non-constant on C. But first we observe that I contains two distinct elements, one in $\{i_1, i_2\}$ and the other in $\{i_3, i_4\}$; indeed otherwise (31) would force C to be contain in a special subvariety of positive codimension. Let $C' \subset Y(1)^{\#I}$ be the Zariski closure of the projection of C onto the coordinates in I. We have $\dim C' = 1$.

Let us treat the case $I = \{i_1, i_2, i_3, i_4\}$; none of the coordinate functions in question is constant on C. The union of the $Y_{\alpha,\beta}$ over the complexification of the semi-algebraic curve from above about some non-singular point is in a hypersurface of $\mathbb{H}^n$ that intersects Z in uncountably many points. The projection of this hypersurface to the coordinates in I is a hypersurface of $\mathbb{H}^4$. Moreover, it meets uncountable many points of $j^{-1}(C'(\mathbb{C}))$. Furthermore, by an argument as in the beginning of this proof we see it contains a neighborhood in $j^{-1}(C'(\mathbb{C}))$. By Proposition 3.1

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

we conclude that C' is in a geodesic subvariety of positive codimension. This is impossible since no coordinate function is constant on C' and since C is not in a special subvariety of positive codimension.

Note that this would not lead to a contradiction if the $Y_{\alpha,\beta}$ were hypersurfaces, i.e. of codimension one rather than two, for then the union over a one-parameter complex family could be all of $\mathbb{H}^n$. It is here where we need that the points in Theorem 1 are contained in a special subvariety of codimension at least two.

Now let us consider the case $\#I = 3$ and $i_3 \notin I$, say. The set of points $\{(z, \alpha z, \beta z_{i_3})\}$ where (α, β) runs over the semi-algebraic curve given above and z runs over $\mathbb{H}$ is contained in a hypersurface of $\mathbb{H}^3$. This hypersurface contains uncountably many points of $j^{-1}(C'(\mathbb{C}))$ because the i_3rd coordinate is constant on C. Therefore, the hypersurface contains a neighborhood in $j^{-1}(C'(\mathbb{C}))$. We again apply Proposition 3.1 and arrive at a contradiction. The more general case $\#I = 3$ and $x_{i_j} \notin I$ follows along the same lines.

In a similar manner one treats the remaining case $\#I = 2$.

Hence each $f_{z,i}^{(j)}$ must have vanishing differential at all points outside $S_z^{(j)}$. Then all the $f_{z,i}^{(j)}$ are constant on each connected component of $W_z^{(j)} \smallsetminus S_z^{(j)}$, which are regular of highest dimension at every point. The corresponding intersections with Z are constant and account for at most $c_6 L^{\delta/2}$ points, i.e. at most a uniformly bounded constant times the number of blocks. However, we have at least $c_7 L^{\delta}$ intersection points in Z to account for.

Therefore, there is a set $S_z^{(j)}$ with at least $c_8 L^{\delta}$ rational points up to height $c_2 L^{c_3}$, corresponding to at least $c_9 L^{\delta}$ distinct intersection points in Z for all sufficiently large L. We repeat the application of Theorem 5 to obtain block families containing the rational points for the fibers in the families $S^{(j)}$. The dimensions of the sets are decreasing. Hence if L is sufficiently large in terms of the constants involved in the induction, they depend only on C, we must eventually find a non-constant function f parameterizing the intersection points with Z. Otherwise we would contradict the lower bound for the number of intersection points in Z. However, that leads to a contradiction of Proposition 3.1 as above. Therefore, L is bounded in terms of C. □

5.3 One special coordinate and one linked pair

Observe that in the following proposition we do not assume that C is asymmetric.

Proposition 5.2. *Let $C \subset Y(1)^n$ be an irreducible curve defined over $\overline{\mathbb{Q}}$ that is not contained in any special subvariety of positive codimension. Let $i_1, i_2, i_3 \in \{1, \dots, n\}$ be pairwise distinct. There are only finitely many $(x_1, \dots, x_n) \in C(\overline{\mathbb{Q}})$ with x_{i_1} special and for which there is a positive integer N with $\Phi_N(x_2, x_3) = 0$.*

Proof. The argument is similar, with just a slight elaboration for the special coordinate. Let $\delta = 1/6$ and let $c_1, c_2, \dots$ denote positive constants that depend on C only, unrelated to previous choices for these constants. We take the definable sets

$$\begin{aligned} Z &= j^{-1}(C(\mathbb{C})) \cap F^n, \\ Y_{a,\alpha} &= \{(z_1, \dots, z_n) \in \mathbb{H}^n;\ z_{i_1} = a,\ z_{i_3} = \alpha z_{i_2}\}, \\ Y &= \{(z_1, \dots, z_n, a, \alpha) \in \mathbb{H}^n \times \mathbb{H} \times \mathrm{GL}_2(\mathbb{R})^+;\ z_{i_1} = a,\ z_{i_3} = \alpha z_{i_2}\}, \\ X &= \{(a, \alpha) \in \mathbb{H} \times \mathrm{GL}_2(\mathbb{R})^+;\ Y_{a,\alpha} \cap Z \neq \emptyset\}. \end{aligned}$$

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

Suppose $x = (x_1, \dots, x_n)$ is as in the hypothesis. A good notion of complexity for the special subvariety containing x is $L = \max\{|\Delta(x_1)|, N\}$. According to Lemma 4.4, the point $(x_1, \dots, x_n)$ has at least $c_1 L^\delta$ conjugates $(x'_1, \dots, x'_n)$ over $\mathbb{Q}$, of which some fixed proportion lie again on C. This gives rise, via Lemma 5.2 for pre-images in F of (x'_{i_2}, x'_{i_3}) and Lemma 5.3 for the pre-images in F of x'_{i_1}, to at least $c_2 L^\delta$ points $(a, \alpha) \in X$. By Lemma 5.3 real and imaginary parts of a are quadratic of height at most $2L$, and α is rational of height at most $c_3 L^{c_4}$. We now apply Theorem 5 to the quadratic points on X to get our block families, and complete the argument as in the proof of Proposition 5.1. □

5.4 Proof of Theorems 1 and 2

Proof of Theorem 1. A special subvariety of codimension at least two is contained in a special subvariety of codimension two. The latter is defined by two equations, each one of which either asserts that some coordinate is special, or links two coordinates by a modular polynomial. Because C is assumed to be defined over $\overline{\mathbb{Q}}$ and since its intersection with any special subvariety of positive codimension is finite we have $C(\mathbb{C}) \cap \mathcal{S}^{[2]} \subset C(\overline{\mathbb{Q}})$. The various cases that may occur are dealt with by the preceding propositions.

Specifically, for any choice of two distinct pairs of distinct indices $\{i_1, i_2\}, \{i_3, i_4\}$, the number of points $(x_1, \dots, x_n)$ satisfying $\Phi_M(x_{i_1}, x_{i_2}) = 0, \Phi_N(x_{i_3}, x_{i_4}) = 0$ for some M, N is finite by Proposition 5.1. The case where one coordinate x_{i_1} is special and two others x_{i_2}, x_{i_3} are linked by a modular relation is dealt with in Proposition 5.2. The case where two coordinates x_{i_1}, x_{i_2} are special reduces to the André–Oort conjecture, as observed in Lemma 4.1. This completes the proof of Theorem 1. □

Proof of Theorem 2. The proof proceeds as for Theorem 1, but since we restrict attention to special but not strongly special subvarieties of codimension at least two we need not consider the case of two pairs of coordinates linked by modular relations. The remaining cases do not require C to be asymmetric. This completes the proof of Theorem 2. □

6. The Mordell conjecture in $Y(1)^n$

Before commencing the proof of Theorem 3, we observe a special case of Lemma 4.2. Suppose $u \in Y(1)(\overline{\mathbb{Q}})$ is non-special. There is a constant $c(u) > 0$ such that if $x \in \overline{\mathbb{Q}}$ and $N \geqslant 1$ with $\Phi_N(x, u) = 0$ then $[\mathbb{Q}(x) : \mathbb{Q}] \geqslant c(u) N^{1/6}$. This follows from Lemma 4.2 by considering the curve C in $Y(1)^2$ whose second coordinate is constant u and $\{i_1, i_2\} = \{i_3, i_4\} = \{1, 2\}$. This lower bound may also be proved by a direct application of the Masser–Wüstholz isogeny estimates in the form proved by Pellarin or by using Serre's open image theorem [Ser72].

Proof of Theorem 3. As usual, definable means definable in $\mathbb{R}_{\mathrm{an,exp}}$ if not stated otherwise.

Since special points and the u for $(i, u) \in U$ are in $\overline{\mathbb{Q}}$, we may assume that V is defined over $\overline{\mathbb{Q}}$. We use the lower bound for Galois orbits observed above combined with the methods of [Pil11, Pil09a]. In the following, $\delta = 1/6$ while $c_1, c_2, \dots$ denote positive constants that may depend at most on C and U.

Consider first the case that U is a set of the form $\{(1, u_1), \dots, (n, u_n)\}$ with each u_i non-special. Set $u = (u_1, \dots, u_n)$. A point $x = (x_1, \dots, x_n) \in Y(1)^n(\overline{\mathbb{Q}})$ is called u-special if, for each coordinate x_i, there is an N_i such that $\Phi_{N_i}(x_i, u_i) = 0$. Since the u_i are non-special the N_i are unique. That is, for the moment we exclude special points and allow just one Hecke orbit for each coordinate. A geodesic subvariety in $Y(1)^n$ is called u-special if it contains a u-special point.

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

We combine the uniformization

$$j : \mathbb{H}^n \to \mathbb{C}^n$$

with the semi-algebraic map

$$\theta : (\mathrm{GL}_2(\mathbb{R})^+)^n \to \mathbb{H}^n, \quad (g_1, \dots, g_n) \mapsto (g_1\sigma_1, \dots, g_n\sigma_n),$$

where the σ_i are j-pre-images of the u_i. We can take $\sigma_i \in F$ (although up to some fixed $\mathrm{SL}_2(\mathbb{Z})$ transformations this makes no difference). A famous result of Schneider states that $z \in \mathbb{H}$ and $j(z)$ are both algebraic only if z is imaginary quadratic, cf. [Bak75, Theorem 6.3]. Since the u_i are algebraic but not special, the σ_i are transcendental, but the rationality will now come from the coordinates in $(\mathrm{GL}_2(\mathbb{R})^+)^n$. A pre-image of a u-special point gives rise to a rational pre-image under θ. The fact that there are multiple (even an algebraic family of) pre-images does not matter since all the rational points up to a specified height are contained in the blocks afforded by Theorem 5.

We have the definable set $Z = j^{-1}(V(\mathbb{C})) \cap F^n$. The map θ is definable, so

$$X = \theta^{-1}(Z)$$

is a definable set. For a u-special point $x \in Y(1)^n(\overline{\mathbb{Q}})$, we use the complexity

$$L = L(x) = \max\{N_1, \dots, N_n\}.$$

Since V is defined over $\overline{\mathbb{Q}}$ and by the observations made at the beginning of this section, if $x \in V(\overline{\mathbb{Q}})$ then x has at least $c_1 L^\delta$ conjugates x' also on V.

Each conjugate has a pre-image in Z, and by Lemma 5.2 it has a pre-image of the form $(g_1\sigma_1, \dots, g_n\sigma_n)$ where $g_i \in \mathrm{GL}_2(\mathbb{Q})^+$ have height at most $c_2 L^{c_3}$ for some absolute c_3. This gives rise to at least

$$c_1 L^\delta$$

rational points of height at most $c_2 L^{c_3}$ on X, corresponding to at least $c_1 L^\delta$ different points in Z.

We apply Theorem 5 to X with $\epsilon = \delta/(2c_3)$. The rational points up to height $c_2 L^{c_3}$ are contained in at most

$$c_4 L^{\delta/2}$$

blocks. The images of a block under θ is a finite union of a uniformly bounded number of blocks in Z, and these blocks in Z contain the pre-image points of the $x^{(i)}$ in Z.

By [Pil11, Theorem 6.8], the maximal algebraic subvarieties of $j^{-1}(V)$ are geodesic. In particular there can be at most

$$c_4 L^{\delta/2}$$

blocks that reduce to points, and so at most that number of points that do not lie on some positive dimensional geodesic subvariety, which we may take to be u-special (the non-u-special ones do not contain any u-special points).

The incompatibility of the upper and lower bounds gives the conclusion that L is bounded if we assume that there are no u-special subvarieties of positive dimension contained in V. The rest of the proof is the same as the corresponding parts of [Pil11, §§ 10 and 11]. By [Pil11, Proposition 10.2], only finitely many geodesic subvarieties occur, up to 'translation', meaning the choice of any fixed coordinates. For each such geodesic subvariety, the u-special translations of it contained in V correspond to special points on some lower-dimensional 'quotient' variety. The proof that V contains only finitely many maximal u-special subvarieties is completed as in [Pil11, Theorem 11.2, § 11.3] by an induction.

https://doi.org/10.1112/S0010437X11005604 Published online by Cambridge University Press

Let us now consider a more general situation. Let S be a subset of $\{1,\dots,n\}$, possibly empty, and $u : S \to \overline{\mathbb{Q}}$ a map. A point $(x_1,\dots,x_n) \in Y(1)^n(\mathbb{C})$ is called u-special if x_i is in the Hecke orbit of $u(i)$ for each $i \in S$, and x_i is special for all other i. A u-special subvariety of $Y(1)^n$ is a u-special point or a geodesic subvariety of positive dimension that contains at least one u-special point.

Permute the coordinates of $Y(1)^n$ so that $S = \{1,\dots,m\}$. We may assume $m > 0$ or the assertion is just André–Oort. We take the uniformization

$$(\mathrm{GL}_2(\mathbb{R})^+)^m \times \mathbb{H}^{n-m} \to \mathbb{H}^n \to Y(1)^n$$

where the left map on the $\mathrm{GL}_2(\mathbb{R})^+$ factors sends g_i to $g_i\sigma_i$, and is the identity on the remaining factors, and the right map is j. The left map is semi-algebraic and therefore definable. Let $Z = j^{-1}(V) \cap F^n$, definable, and X be its pre-image in $(\mathrm{GL}_2(\mathbb{R})^+)^m \times \mathbb{H}^{n-m}$. Then X is definable, and u-special points in V give rise to points in X that are rational in the first n factors and quadratic in the other factors. The proof in this case proceeds combining the proof above with the proof in [Pil11].

By an obvious combinatorial argument one can allow each coordinate to come from a finite set of Hecke orbits rather than a single Hecke orbit. This completes the proof of Theorem 3. □

We remark that, following [Pil11], we can replace the factors $Y(1)$ in Theorem 3 by modular curves $Y(\Gamma_i) = \Gamma_i\backslash\mathbb{H}$ for any congruence subgroups $\Gamma_1,\dots,\Gamma_n$ of $\mathrm{SL}_2(\mathbb{Z})$, and we can further, as in [Pil11], combine the statement with the Manin–Mumford conjecture for abelian varieties over $\overline{\mathbb{Q}}$ or with Manin–Mumford for products of linear tori and elliptic curves defined over $\overline{\mathbb{Q}}$. It would be interesting to see if one could allow finite generation also in the other factors and establish the 'Mordell–Lang' statement for such X.

Acknowledgements

We are very grateful to Daniel Bertrand for helpful conversations in the course of our work, especially on the aspects involving functional transcendence properties of the j-function. Both authors would also like to thank Umberto Zannier for fruitful discussions, and the organizers of the symposium on new directions in the model theory of fields in Durham 2009 for providing an opportunity for model-theorists and number-theorists to meet. Our collaboration began there. Habegger thanks Jochen Koenigsmann and the Mathematical Institute, Oxford, for the invitation and for the stimulating atmosphere. He also thanks Lars Kühne and Gisbert Wüstholz for discussions related to the former's Master Thesis on the André–Oort conjecture. Habegger was supported by an ETH Fellowship grant. Pila thanks Roger Heath-Brown and the Mathematical Institute, Oxford, for having him as an academic visitor in the period in which this research was done, and he is grateful to the Leverhulme Trust for supporting his work with a Research Fellowship.

References

And92 Y. André, *Mumford–Tate groups of mixed Hodge structures and the theorem of the fixed part*, Compositio Math. **82** (1992), 1–24.

And98 Y. André, *Finitude des couples d'invariants modulaires singuliers sur une courbe algébrique plane non modulaire*, J. Reine Angew. Math. **505** (1998), 203–208.

Ax71 J. Ax, *On Schanuel's conjectures*, Ann. of Math. (2) **93** (1971), 252–268.

Bak75 A. Baker, *Transcendental number theory* (Cambridge University Press, Cambridge, 1975).

BG06 E. Bombieri and W. Gubler, *Heights in diophantine geometry* (Cambridge University Press, Cambridge, 2006).

BMZ08 E. Bombieri, D. Masser and U. Zannier, *On unlikely intersections of complex varieties with tori*, Acta Arith. **133** (2008), 309–323.

Bre01 F. Breuer, *Heights of CM points on complex affine curves*, Ramanujan J. **5** (2001), 311–317.

Cas57 J. W. S. Cassels, *An introduction to diophantine approximation* (Cambridge University Press, Cambridge, 1957).

Cox89 D. A. Cox, *Primes of the form* $x^2 + ny^2$ (John Wiley & Sons, New York, 1989).

Dav00 H. Davenport, *Multiplicative number theory* (Springer, Berlin, 2000).

Dav95 S. David, *Minorations de formes linéaires de logarithmes elliptiques*, Mém. Soc. Math. Fr. (N.S.) **62** (1995), 1–143.

Del71 P. Deligne, *Théorie de Hodge: II*, Publ. Math. Inst. Hautes Études Sci. **40** (1971), 5–57.

Dri84 L. van den Dries, *Remarks on Tarski's problem concerning* $(\mathbf{R}, +, \cdot, \exp)$, in *Logic colloquium '82 (Florence, 1982)*, Studies in Logic and the Foundations of Mathematics, vol. 112 (North-Holland, Amsterdam, 1984), 97–121.

Dri86 L. van den Dries, *A generalization of the Tarski–Seidenberg theorem, and some nondefinability results*, Bull. Amer. Math. Soc. (N.S.) **15** (1986), 189–193.

Dri98 L. van den Dries, *Tame topology and o-minimal structures*, London Mathematical Society Lecture Note Series, vol. 248 (Cambridge University Press, Cambridge, 1998).

DMM94 L. van den Dries, A. Macintyre and D. Marker, *The elementary theory of restricted analytic fields with exponentiation*, Ann. of Math. (2) **140** (1994), 183–205.

DM94 L. van den Dries and C. Miller, *On the real exponential field with restricted analytic functions*, Israel J. Math. **85** (1994), 19–56.

DM96 L. van den Dries and C. Miller, *Geometric categories and o-minimal structures*, Duke Math. J. **84** (1996), 497–540.

Edi98 B. Edixhoven, *Special points on the product of two modular curves*, Compositio Math. **114** (1998), 315–328.

Edi05 B. Edixhoven, *Special points on products of modular curves*, Duke Math. J. **126** (2005), 325–348.

EMOT81 A. Erdélyi, W. Magnus, F. Oberhettinger and F. G. Tricomi, *Higher transcendental functions, Vol. I* (Robert E. Krieger, Melbourne, FL, 1981).

Gab68 A. M. Gabrièlov, *Projections of semianalytic sets*, Funktsional. Anal. i Prilozhen. **2** (1968), 18–30.

GS00 A. Granville and H. M. Stark, *ABC implies no 'Siegel zeros' for L-functions of characters with negative discriminant*, Invent. Math. **139** (2000), 509–523.

GR84 H. Grauert and R. Remmert, *Coherent analytic sheaves* (Springer, Berlin, 1984).

Hab07 P. Habegger, *Heights and multiplicative relations on algebraic varieties*, PhD thesis, University of Basel (2007).

Hab10 P. Habegger, *Weakly bounded height on modular curves*, Acta Math. Vietnam. **35** (2010), 43–69.

Hus04 D. Husemöller, *Elliptic curves* (Springer, Berlin, 2004).

Ima76 H. Imai, *On the Hodge groups of some abelian varieties*, Kōdai Math. Semin. Rep. **27** (1976), 367–372.

Kol68 E. R. Kolchin, *Algebraic groups and algebraic dependence*, Amer. J. Math. **90** (1968), 1151–1164.

Lan87 S. Lang, *Elliptic functions* (Springer, Berlin, 1987).

MW90 D. W. Masser and G. Wüstholz, *Estimating isogenies on elliptic curves*, Invent. Math. **100** (1990), 1–24.

Mau08 G. Maurin, *Courbes algébriques et équations multiplicatives*, Math. Ann. **341** (2008), 789–824.

Moo98 B. Moonen, *Linearity properties of Shimura varieties. I*, J. Algebraic Geom. **7** (1998), 539–567.

NT91 Y. Nakkajima and Y. Taguchi, *A generalization of the Chowla–Selberg formula*, J. Reine Angew. Math. **419** (1991), 119–124.

Pel01 F. Pellarin, *Sur une majoration explicite pour un degré d'isogénie liant deux courbes elliptiques*, Acta Arith. **100** (2001), 203–243.

PS04 Y. Peterzil and S. Starchenko, *Uniform definability of the Weierstrass $\wp$ functions and generalized tori of dimension one*, Selecta Math. (N.S.) **10** (2004), 525–550.

Pil09a J. Pila, *On the algebraic points of a definable set*, Selecta Math. (N.S.) **15** (2009), 151–170.

Pil09b J. Pila, *Rational points of definable sets and results of André–Oort–Manin–Mumford type*, Int. Math. Res. Not. IMRN (2009), 2476–2507.

Pil11 J. Pila, *O-minimality and the André–Oort conjecture for $\mathbb{C}^n$*, Ann. of Math. (2) **173** (2011), 1779–1840.

PW06 J. Pila and A. J. Wilkie, *The rational points of a definable set*, Duke Math. J. **133** (2006), 591–616.

PZ08 J. Pila and U. Zannier, *Rational points in periodic analytic sets and the Manin–Mumford conjecture*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **19** (2008), 149–162.

PS86 A. Pillay and C. Steinhorn, *Definable sets in ordered structures. I*, Trans. Amer. Math. Soc. **295** (1986), 565–592.

Pin05 R. Pink, *A common generalization of the conjectures of André–Oort, Manin–Mumford, and Mordell–Lang*, Preprint (2005), available at www.math.ethz.ch/∼pink/ftp/ (AOMMML.pdf).

Ray85 M. Raynaud, *Hauteurs et isogénies*, in *Seminar on arithmetic bundles: the Mordell conjecture* (Paris, 1983/84), Astérisque **127** (1985), 199–234.

Ser72 J.-P. Serre, *Propriétés galoisiennes des points d'ordre fini des courbes elliptiques*, Invent. Math. **15** (1972), 259–331.

UY09 E. Ullmo and A. Yafaev, *The André–Oort conjecture for products of modular curves*, in *Arithmetic geometry*, Clay Mathematics Proceedings, vol. 8 (American Mathematical Society, Providence, RI, 2009), 431–439.

Wil96 A. J. Wilkie, *Model completeness results for expansions of the ordered field of real numbers by restricted Pfaffian functions and the exponential function*, J. Amer. Math. Soc. **9** (1996), 1051–1094.

Zil02 B. Zilber, *Exponential sums equations and the Schanuel conjecture*, J. Lond. Math. Soc. (2) **65** (2002), 27–44.

P. Habegger habegger@math.uni-frankfurt.ch
Institut fuer Mathematik, Goethe Universitaet Frankfurt, Robert-Mayer-Strasse 6-8,
60325 Frankfurt am Main, Germany

J. Pila pila@maths.ox.ac.uk
School of Mathematics, University of Bristol, Bristol, BS8 1TW, UK

and (current address): Mathematical Institute, University of Oxford, 24-29 St Giles',
Oxford, OX1 3LB, UK