[O-minimalité et certaines intersections atypiques]

O-MINIMALITY AND CERTAIN ATYPICAL INTERSECTIONS

P. HABEGGER AND J. PILA

Abstract. We show that the strategy of point counting in o-minimal structures can be applied to various problems on unlikely intersections that go beyond the conjectures of Manin-Mumford and André-Oort. We verify the so-called Zilber-Pink Conjecture in a product of modular curves on assuming a lower bound for Galois orbits and a sufficiently strong modular Ax-Schanuel Conjecture. In the context of abelian varieties we obtain the Zilber-Pink Conjecture for curves unconditionally when everything is defined over a number field. For higher dimensional subvarieties of abelian varieties we obtain some weaker results and some conditional results.

On démontre que la stratégie de comptage dans des structures o-minimales est suffisante pour traiter plusieurs problèmes qui vont au-delà des conjectures de Manin-Mumford et André-Oort. On vérifie la conjecture de Zilber-Pink pour un produit de courbes modulaires en supposant une minoration assez forte pour la taille de l'orbite de Galois et en supposant une version modulaire du théorème de Ax-Schanuel. Dans le cas des variétés abéliennes on démontre la conjecture de Zilber-Pink pour les courbes si tous les objets sont définis sur un corps de nombres. Pour les sous-variétés de dimension supérieure on obtient quelques résultats plus faibles et quelques résultats conditionnels.

Key words: Zilber-Pink conjecture, unlikely intersections, o-minimality.
Mot clés: Conjecture de Zilber-Pink, intersections exceptionnelles, o-minimalité.

Mathematics Subject Classification (2010): 11G10, 11G18 (primary), 03C64, 11G50, 14K12, 14G35, 14G40 (secondary).

Contents

1. Introduction

The object of this paper is to show that the "o-minimality and point-counting" strategy can be applied to quite general problems of "unlikely intersection" type as formulated in the Zilber-Pink Conjecture (ZP; see Section 2 for various formulations), provided one assumes certain arithmetic and functional transcendence hypotheses. In these problems there is an ambient variety X of a certain type equipped with a distinguished collection $\mathcal{S}$ of "special" subvarieties. The conjecture governs the intersections of a subvariety $V \subseteq X$ with the members of $\mathcal{S}$. In the problems we consider, X will be either a product of non-compact modular curves (for which it is sufficient to consider the case $X = Y(1)^n$) or an abelian variety, but the same formulations should be applicable more generally. In this paper, a subvariety is always geometrically irreducible and therefore in particular non-empty. A curve is a subvariety of dimension 1.

Our most general results are conditional, but let us state first an unconditional result in the abelian setting.

Say X is an abelian variety defined over a field K and $\overline{K}$ is a fixed algebraic closure of K. For any $r \in \mathbb{R}$ we write

$$X^{[r]} = \bigcup_{\operatorname{codim}_X H \geq r} H(\overline{K})$$

where H runs over algebraic subgroups of X satisfying the dimension condition.

Theorem 1.1. *Let X be an abelian variety defined over a number field K and suppose $V \subseteq X$ is a curve, also defined over K. If V is not contained in a proper algebraic subgroup of X, then $V(\overline{K}) \cap X^{[2]}$ is finite.*

This theorem is the abelian version of Maurin's Theorem [29]. We will see a more precise version in Theorem 9.14.

We briefly describe previously known cases of Theorem 1.1 under additional hypotheses on V or X. Viada [49] proved finiteness for V not contained in the translate of a proper abelian subvariety and if X is the power of an elliptic curve with complex multiplication. Rémond and Viada [45] then removed the hypothesis on V. This was later generalised by Ratazzi [40] to when the ambient group variety is isogenous to a power of an abelian variety with complex multiplication. Carrizosa's height lower bound [12, 13] in combination with Rémond's height upper bound [43] led to a proof for all abelian varieties with complex multiplication. Work of Galateau [18] and Viada [50] cover the case of an arbitrary product of elliptic curves.

More generally we show, in the abelian and modular settings, that the Zilber-Pink conjecture may be reduced to two statements, one of an arithmetic nature, the other a functional transcendence statement. In general, the former statement remains conjectural in both settings. In the abelian setting, the functional transcendence statement follows from a theorem of Ax [3], while in the modular setting a proof of it has been announced recently by Pila-Tsimerman [34]. Both statements are generalisations of statements which have been used to establish cases of the André-Oort conjecture, and this aspect of our work is in the spirit of Ullmo [47].

The arithmetic hypothesis, which we formulate here, is the "Large Galois Orbit" hypothesis (LGO) and asserts that, for fixed $V \subseteq X$, certain ("optimal") isolated intersection points of V with a special subvariety T have a "large" Galois orbit over a fixed finitely generated field of definition for V, expressed in terms of a suitable complexity measure of T. Special subvarieties in our settings are described in Section 2 for abelian varieties and Section 3.2 for $Y(1)^n$ and LGO is formulated in Section 8.

In the context of the André-Oort Conjecture, the Generalised Riemann Hypothesis (GRH) suffices to guarantee LGO (see [46, 48]). However, it is not clear to the authors if a variant of the Riemann Hypothesis leads to large Galois orbits for isolated points in $V \cap T$ if $\dim T \geq 1$. Indeed, in the Shimura setting, there seems to be no Galois-theoretic description of isolated points in $V \cap T$ which is rooted in class field theory. On the other hand, suitable bounds have been established unconditionally for André-Oort in several cases, and perhaps LGO will be found accessible without assuming GRH.

Associated with X is a certain transcendental uniformisation $\pi : U \to X$. The functional transcendence hypothesis is the "Weak Complex Ax" hypothesis (WCA) and is a weak form of an analogue for π of "Ax-Schanuel" for cartesian powers of the exponential function. The latter result, due to Ax [2], affirms Schanuel's Conjecture (see [24, p. 30]) for differential fields. WCA is formulated in Section 5.

In the modular case $X = Y(1)^n$ our result is the following. A very special case of it was established unconditionally by us in [21].

Theorem 1.2. *If LGO and WCA hold for $Y(1)^n$ then the Zilber-Pink Conjecture holds for subvarieties of $Y(1)^n$ defined over $\mathbb{C}$. Moreover, if WCA holds for $Y(1)^n$ and LGO holds with $K = \mathbb{Q}$, then the Zilber-Pink Conjecture holds for subvarieties of $Y(1)^n$ defined over $\overline{\mathbb{Q}}$.*

In the case that X is an abelian variety, we establish the same result in Section 9. However, as mentioned above, in this case WCA is known, and LGO can be established in the case that V is one-dimensional when X and V are defined over $\overline{\mathbb{Q}}$. This allows us to prove the above unconditional result for curves.

All current approaches towards Theorem 1.1 require a height upper bound on the set of points in question. Like many of the papers cited above we use Rémond's height bound [43] which relies on his generalisation of Vojta's height inequality.

In contrast to previous approaches we do not rely on delicate Dobrowolski-type [12, 13, 40] or Bogomolov-type [18] height lower bounds to pass from bounded height to finiteness. These height lower bounds are expected (but not known) to generalise to arbitrary abelian varieties. Instead we will use a variation of the strategy originally devised by Zannier to reprove the Manin-Mumford Conjecture [36] for abelian varieties. This approach relied on the Pila-Wilkie point counting result in o-minimal structures. We will still require a height lower bound. However, the robust nature of the method allows us to use Masser's general bound [27] which predates the sophisticated and essentially best-possible results of Ratazzi and Carrizosa that require the ambient abelian variety to have complex multiplication.

In her recent Ph.D. thesis, Capuano [11] counted rational points on suitable definable subsets of a Grassmanian to obtain finiteness results on unlikely intersections with curves in the algebraic torus.

In the next theorem we collect several partial results in the abelian setting for subvarieties of arbitrary dimension.

Theorem 1.3. *Let $V \subseteq X$ be a subvariety of an abelian variety, both defined over a number field K. Let us also fix an ample, symmetric line bundle on X and its associated Néron-Tate height $\hat{h}$.*

(i) *If $S \geq 0$ then*

$$\left\{P \in V(\overline{K}) \cap X^{[1+\dim V]};\ \hat{h}(P) \leq S\right\}$$

is contained in a finite union of proper algebraic subgroups of X.

(ii) *Suppose $\dim \varphi(V) = \min\{\dim X/Y, \dim V\}$ for all abelian subvarieties $Y \subseteq X$ where $\varphi : X \to X/Y$ is the canonical morphism. If $\dim V \geq 1$ then $V(\overline{K}) \cap X^{[1+\dim V]}$ is not Zariski dense in V.*

Theorem 9.15 refines this statement. A particularly simple example of a surface that does not satisfy the hypothesis in (ii) is the square $V \times V$ of a curve $V \subsetneq X$. The Zilber-Pink Conjecture remains open for surfaces in abelian varieties defined over a number field. In Corollary 9.9 we show that a sufficiently strong height upper bound leads to a proof of the Zilber-Pink Conjecture for abelian varieties over number fields.

The burden of the theorems is that the o-minimality/point-counting strategy is adequate to deal with “atypical intersections”, given these additional ingredients. To some extent this is already demonstrated for curves by the work of Masser and Zannier [28] and the authors’ earlier work [21]. Here we will handle subvarieties of arbitrary dimension and confirm that the o-minimal method scales to this generality. Ullmo [47] shows that a “large Galois orbit” statement, “Ax-Lindemann”, and a height upper bound for certain pre-images of special points, together with a suitable definability result, enable a proof of the André-Oort Conjecture by point-counting and o-minimality. In our setting, the special subvarieties generally have positive dimension and are often unaffected by the Galois action. Rather we must count objects that arise when intersecting them with the given subvariety. More generally, one can formulate results along the lines of the Counting Theorem of [35] for “atypical intersections” of a definable set in an o-minimal structure with linear subvarieties defined over $\mathbb{Q}$ (or indeed the members of any definable family of subvarieties having rational parameters). Here (as in previous results by these methods) o-minimality is used for more than point-counting.

For the basics about o-minimality see [15]. The definability properties required in this paper are afforded by the result, due to Peterzil-Starchenko [31], that the j-function restricted to the usual fundamental domain $\mathbb{F}_0$ for the $\mathrm{SL}_2(\mathbb{Z})$ action on the upper half plane $\mathbb{H}$ is definable in the o-minimal structure $\mathbb{R}_{\mathrm{an,exp}}$ (see [16]). Accordingly, in this paper “definable” will mean “definable in $\mathbb{R}_{\mathrm{an,exp}}$” unless stated otherwise. However, the exponential function is superfluous when working with abelian varieties. Here it is enough to work in $\mathbb{R}_{\mathrm{an}}$, the structure generated by restricted real analytic functions, which was recognised as being o-minimal by van den Dries after work of Gabrielov. In Section 7 we will work with more general o-minimal structures.

The reader who is mainly interested in Theorem 1.1 on curves in abelian varieties can skip over several sections in this paper. Indeed, many steps required in the higher dimensional case simplify considerably. So we briefly indicate a minimalist approach to Theorem 1.1. A sketch of how our proof of the more general Theorem 9.14 plays out in the case of curves is given at the end of Section 9. A good starting point is the description of basic properties of algebraic subgroups of an abelian variety in Section 3.1. Next, Theorem 5.4 is a version of Ax's Theorem which is suffiently strong for our purposes; it is presented in Section 5.1. Corollary 7.2 of Section 7 is a counting result in the spirit of the Theorem of Pila-Wilkie. We count points on definable sets were a certain collection of coordinates is rational and of bounded height and the remaining coordinates are unrestricted. These ingredients are then mixed in Section 9 together with a height lower bound of Masser and a height upper bound of Rémond to get Theorem 9.14, a stronger version of Theorem 1.1.

2. The Zilber-Pink Conjecture

The Zilber-Pink Conjecture is a far-reaching generalisation of the Mordell-Lang and André-Oort Conjectures. Different formulations in different settings, but based on the same underlying idea of "atypical" or "unlikely" intersections, were made by Zilber [53], Pink [37, 38], and Bombieri-Masser-Zannier [10]. We will not address aspects of uniformity over families of subvarieties here.

Bombieri-Masser-Zannier [8] show that all the versions are equivalent for $\mathbb{G}_m^n$. Essentially the same argument shows they are equivalent for $Y(1)^n$, but in the general case this is unclear. The version we give is in any case the strongest: essentially it is the Zilber or Bombieri-Masser-Zannier statement in Pink's general setting, where the ambient variety is a mixed Shimura variety (see [38]). For an introduction to the conjecture and the state-of-the-the-art see [52].

The general setting involves an ambient variety which is a *mixed Shimura variety* (see [37]). A mixed Shimura variety X is endowed with a (countable) collection $\mathcal{S} = \mathcal{S}_X$ of *special subvarieties* and a larger (usually uncountable) collection $\mathcal{W} = \mathcal{W}_X$ of *weakly special* subvarieties. Special subvarieties of dimension zero are called *special points.* We do not provide a general definition for such subvarieties but rather refer to Pink's paper [37] for details. This paper is only concerned with the cases when X is an abelian variety or the product $Y(1)^n$ of the modular curve. In the former case, the special subvarieties are defined below. A detailed description of special subvarieties in the modular setting is given in section 3.2.

Say X is an algebraic torus $\mathbb{G}_m^n$, an abelian variety, or even a semi-abelian variety defined over $\mathbb{C}$. In general, such X is not a mixed Shimura variety, but rather appears as a weakly special subvariety of a family of semi-abelian varieties, and shares enough formal properties with mixed Shimura varieties to formulate a Zilber-Pink Conjecture on unlikely intersections. A **weakly special subvariety** of X is a **coset**, that is the translate of a connected algebraic subgroup of X. A **special subvariety** of X is a **torsion coset**, that is a coset of X that contains a point of finite order. In particular a **special point** is a **torsion point**, that is a point of finite order in the group. We thus write $\mathcal{W}_X$ and $\mathcal{S}_X$ for the set of all cosets and torsion cosets of X, respectively. We study torsion cosets in more detail in Section 3.1.

Definition 2.1. *Let X be a mixed Shimura variety or a semi-abelian variety defined over $\mathbb{C}$. Let $V \subseteq X$ be a subvariety. A subvariety $A \subseteq V$ is called* **atypical** *(for V in X) if there is a special subvariety T of X such that A is an irreducible component of $V \cap T$ and*

$$\dim A > \dim V + \dim T - \dim X$$

i.e. A is an "atypical component" in Zilber's terminology [53]. The **atypical set** *of V (in X) is the union of all atypical subvarieties, and is denoted* $\mathrm{Atyp}(V, X)$.

A special subvariety T of a mixed Shimura variety X is itself a mixed Shimura variety, with

$$\mathcal{S}_T = \{A;\ A \text{ is an irreducible component of } T \cap S \text{ for some } S \in \mathcal{S}_X\},$$

$$\mathcal{W}_T = \{A;\ A \text{ is an irreducible component of } T \cap S \text{ for some } S \in \mathcal{W}_X\}.$$

The following is a "CIT" (cf. Conjecture 2 [53]) version of Pink's Conjecture, and is on its face stronger than the statement conjectured by Pink. It is convenient to frame it as a statement about all special subvarieties of a given mixed Shimura variety X.

Conjecture 2.2 (Zilber-Pink (ZP) for X)**.** *Let X be a mixed Shimura variety or a semi-abelian variety defined over $\mathbb{C}$. If T is a special subvariety of X and if $V \subseteq T$ is a subvariety, then $\mathrm{Atyp}(V, T)$ is a finite union of atypical subvarieties. Equivalently, V contains only a finite number of maximal atypical (for V in T) subvarieties.*

Now some subvarieties are more atypical than others. Since the collection of special subvarieties is closed under taking irreducible components of intersections and contains X, given $A \subseteq X$ there is a smallest special subvariety containing A which we denote $\langle A\rangle$. We abbreviate $\langle P\rangle = \langle\{P\}\rangle$ for a singleton $\{P\}$. We define (following Pink [38]) the **defect** of A as

$$\delta(A) = \dim\langle A\rangle - \dim A.$$

Then $A \subseteq V$ is atypical for V in X if $\delta(A) < \dim X - \dim V$.

The atypical set is simply a union of atypical subvarieties of V, and it may happen that an atypical subvariety is contained in some larger but less atypical subvariety. A generalisation of the argument showing that ZP implies the André-Oort Conjecture (which corresponds to subvarieties of defect 0) shows that ZP implies a notionally stronger version in which one considers subvarieties of each defect separately.

For a subvariety $V \subseteq X$ and a non-negative integer δ denote by

$$\mathrm{Atyp}^{\delta}(V)$$

the union of subvarieties $A \subseteq V$ with $\delta(A) \leq \delta$. We observe that $\mathrm{Atyp}^{\delta}(V) = V$ if $\delta \geq \delta(V) = \dim\langle V\rangle - \dim V$; and so this case is of no interest.

If $\delta < \delta(V)$ and if $A \subseteq V$ satisfies $\delta(A) \leq \delta$, then A is contained in an irreducible component of $V \cap \langle A\rangle$, which must be atypical for V in $\langle V\rangle$ and which has defect $\leq \delta$. So for δ in this range the set $\mathrm{Atyp}^{\delta}(V)$ is a union of such atypical subvarieties, and we make the following conjecture (which is trivial for $\delta \geq \delta(V)$).

Conjecture 2.3 (Articulated Zilber-Pink (AZP) for X)**.** *Let X be a mixed Shimura variety or a semi-abelian variety defined over $\mathbb{C}$. Let $V \subseteq X$ be a subvariety and* *let δ be a non-negative integer. Then* $\mathrm{Atyp}^{\delta}(V)$ *is a finite union of subvarieties of* V *of defect* $\leq \delta$.

The following implication uses only the formal properties of special subvarieties; the reverse implication is immediate. A version of this result appears in [53].

Proposition 2.4. *Let X be as in Conjecture 2.3. Then ZP for X implies AZP for X.*

Proof. The proof is by induction on the dimension of the ambient special subvariety $T = \langle V \rangle \subseteq X$. AZP is clear if $\dim T = 0$. So AZP holds for all proper special subvarieties of T. Let $V \subseteq T$. By ZP, there are finitely many atypical subvarieties A_i, with associated special subvarieties $\langle A_i \rangle$, such that every atypical subvariety A of V for T is contained in some A_i. The $\langle A_i \rangle$ are evidently proper ($V \subseteq T$ is not atypical for V in T). Now fix $\delta \geq 0$; we may assume $\delta < \dim\langle V \rangle - \dim V$. Then with $\delta_i = \delta(A_i)$,

$$\mathrm{Atyp}^{\delta}(V) = \left(\bigcup_{i:\delta_i \leq \delta} A_i \right) \cup \left(\bigcup_{i:\delta_i > \delta} \mathrm{Atyp}^{\delta}(A_i) \right)$$

is a finite union by the induction hypothesis, which gives AZP for the A_i. □

Let us formulate one last conjecture in the same spirit as those above. First we require the notion of an "optimal" subvariety which proves to be quite useful in the context of unlikely intersections and which will play an important role in the following sections.

Definition 2.5. *Let X be a mixed Shimura variety or a semi-abelian variety defined over $\mathbb{C}$. Let $V \subseteq X$ be a subvariety. A subvariety $A \subseteq V$ is said to be* **optimal** *(for V in X) if there is no subvariety B with $A \subsetneq B \subseteq V$ such that*

$$\delta(B) \leq \delta(A).$$

The specification of V and X will generally be suppressed, as no confusion should arise. We write $\mathrm{Opt}(V)$ *for the set of all optimal subvarieties for V.*

We will often use the following basic property. It is possible to enlarge a given subvariety $B \subseteq V$ to an optimal subvariety $A \subseteq V$ with $A \supseteq B$ and $\delta(A) \leq \delta(B)$. It is illustrative to consider some more formal properties of an optimal subvariety $A \in \mathrm{Opt}(V)$. Clearly, A is an irreducible component of $V \cap \langle A \rangle$. If $A \neq V$, then $\delta(A) < \delta(V)$ or in other words

$$\dim A > \dim\langle A \rangle + \dim V - \dim\langle V \rangle.$$

So A is atypical for V in $\langle V \rangle$. We will see that the arithmetically interesting case is when $A = \{P\}$ is a singleton. Then P is contained in a special subvariety of dimension strictly less than $\dim\langle V \rangle - \dim V$. The whole subvariety is always optimal, i.e. $V \in \mathrm{Opt}(V)$. So "maximal optimal subvariety" is a useless concept. In a certain sense maximality is built into the notation of optimality. Indeed, if $\delta(A) = 0$, then A is a maximal special subvariety contained completely in X.

Conjecture 2.6. *Let X be a mixed Shimura variety or a semi-abelian variety defined over $\mathbb{C}$ and let $V \subseteq X$ be a subvariety. Then* $\mathrm{Opt}(V)$ *is finite.*

Lemma 2.7. *Let X be a mixed Shimura variety or a semi-abelian variety defined over $\mathbb{C}$. Then the conclusions of Conjectures 2.2 and 2.6 are equivalent for X.*

Proof. Let V be a subvariety of X.

First we suppose that $\mathrm{Opt}(V)$ is finite. Let T be a special subvariety of X containing V. We may assume $\langle V\rangle = T$, as $\mathrm{Atyp}(V,T) = V$ otherwise. Let A be atypical for V in T and let $B \supseteq A$ be an optimal subvariety for V with $\delta(B) \leq \delta(A)$. Then $\delta(B) < \dim T - \dim V$ and so B is also atypical for V in T. Conjecture 2.2 follows for X as B is contained in a member of the finite set $\mathrm{Opt}(V) \smallsetminus \{V\}$ of proper optimal subvarieties for V.

Let us now assume conversely that Conjecture 2.2 holds for V with $T = \langle V\rangle$. We must show that there are only finitely many possibilities for $A \in \mathrm{Opt}(V)$. Clearly, we may assume $A \subsetneq V$. But then $\delta(A) < \delta(V) = \dim T - \dim V$ by optimality. So A is also atypical for V in T. It is contained in a subvariety B that is maximal atypical for V in T. So B comes from a finite set. We observe that $\dim B < \dim V$ and that A lies in $\mathrm{Opt}(B)$, which is finite by induction on the dimension. □

A product of modular curves, in particular $Y(1)^n$, is a (pure) Shimura variety. Another example of a (pure) Shimura variety is the moduli space $\mathcal{A}_g$ of principally polarised abelian varieties of dimension g. It is a special subvariety of a larger mixed Shimura variety $\mathcal{X}_g$ which consists of $\mathcal{A}_g$ fibered at each point by the corresponding abelian variety.

Conjecture 2.2 for an abelian variety X and its special subvarieties is equivalent to ZP as formulated for subvarieties $V \subseteq X \subseteq \mathcal{X}_g$ (see [38, 5.2] where the equivalence is proved for Pink's Conjecture; with obvious modifications the argument proves the equivalence in the version we have given).

3. Special subvarieties

In the next two sections we discuss in more detail the special subvarieties of an abelian variety and of $Y(1)^n$.

3.1. **The abelian setting.**

In the case of abelian varieties, the special subvarieties are the torsion cosets, i.e. the irreducible components of algebraic subgroups or in other words, translates of abelian subvarieties by points of finite order. In this section we recall some basic facts on torsion cosets and we define their "complexity".

An **inner-product** on an $\mathbb{R}$-vector space W is a symmetric, positive definite bilinear form $\langle\cdot,\cdot\rangle : W \times W \to \mathbb{R}$. The **volume** $\mathrm{vol}(\Omega)$ of a finitely generated subgroup Ω of W with respect to $\langle\cdot,\cdot\rangle$ is $\mathrm{vol}(\Omega) = |\det(\langle\omega_i,\omega_j\rangle)|^{1/2}$ for any $\mathbb{Z}$-basis $(\omega_1,\ldots,\omega_\rho)$ of Ω. The volume is independent of the choice of basis. The **orthogonal complement** of a vector subspace $U \subseteq W$ is $U^\perp = \{w \in W;\ \langle w,u\rangle = 0 \text{ for all } u \in U\}$.

Let X be an abelian variety defined over $\mathbb{C}$ with $\dim X = g \geq 1$ and suppose that $\mathcal{L}$ is an ample line bundle on X. The **degree** of X with respect to $\mathcal{L}$ is the intersection number $\deg_{\mathcal{L}} X = (\mathcal{L}^g[X]) \geq 1$.

The line bundle $\mathcal{L}$ defines a hermitian form

$$H : T_0(X) \times T_0(X) \to \mathbb{C}$$

on the tangent space $T_0(X)$ of X at the origin. This form is positive definite since $\mathcal{L}$ is ample. It is $\mathbb{C}$-linear in the first argument and satisfies $H(v,w) = \overline{H(w,v)}$ for $v,w \in T_0(X)$. The real part $\mathrm{Re}(H)$ is an inner-product $\langle\cdot,\cdot\rangle$ on $T_0(X)$ taken as an $\mathbb{R}$-vector space of dimension $2g$. Thus we obtain a norm $\|v\| = \langle v,v\rangle^{1/2}$ on $T_0(X)$. The imaginary part $E = \mathrm{Im}(H)$ is a non-degenerate symplectic form $V \times V \to \mathbb{R}$.

Let $\Omega_X \subseteq T_0(X)$ denote the period lattice of X. It is a free abelian group of rank $2g$ and generates $T_0(X)$ as an $\mathbb{R}$-vector space. Therefore, $\mathrm{vol}(\Omega_X) > 0$ where vol denotes the volume with respect to the inner-product $\langle\cdot,\cdot\rangle_{\mathcal{L}}$. The subgroup Ω_X is discrete in $T_0(X)$.

Lemma 3.1. *We have* $\deg_{\mathcal{L}} X = g!\mathrm{vol}(\Omega_X)$.

Proof. This is a well-known consequence of the Riemann-Roch Theorem for abelian varieties, see Chapter 3.6 [6]. □

Now suppose that $Y \subseteq X$ is an abelian subvariety of dimension $\dim Y \geq 1$. The pull-back $\mathcal{L}|_Y$ of $\mathcal{L}$ by the inclusion map $Y \hookrightarrow X$ is an ample line bundle on Y. We treat $T_0(Y)$ as a vector subspace of $T_0(X)$. The hermitian form on $T_0(Y)$ induced by $\mathcal{L}|_Y$ is just the restriction of H. Let $\Omega_Y \subseteq T_0(Y)$ denote the period lattice of Y. Then $\deg_{\mathcal{L}|_Y} Y = (\dim Y)!\mathrm{vol}(\Omega_Y)$ by the previous lemma. The projection formula implies $\deg_{\mathcal{L}|_Y} Y = (\mathcal{L}^{\dim Y}[Y])$. We will abbreviate this degree by $\deg_{\mathcal{L}} Y$.

The next lemma uses Minkowski's Second Theorem from the Geometry of Numbers.

Lemma 3.2. *There exists a constant $c > 0$ depending only on $(X,\mathcal{L})$ with the following properties.*

(i) *There exist linearly independent periods $\omega_1,\ldots,\omega_{2\dim Y} \in \Omega_Y$ with $\|\omega_i\| \leq c\deg_{\mathcal{L}} Y$ for $1 \leq i \leq 2\dim Y$ and $\|\omega_1\|\cdots\|\omega_{2\dim Y}\| \leq c\deg_{\mathcal{L}} Y$.*

(ii) *If $z \in \Omega_X + T_0(Y)$ there exist $\omega \in \Omega_X$ with $z - \omega \in T_0(Y)$ and $\|\omega\| \leq \|z\| + c\deg_{\mathcal{L}} Y$.*

Proof. Let $0 < \lambda_1 \leq \ldots \leq \lambda_{2\dim Y}$ be the successive minima of Ω_Y with respect to the closed unit ball $\{z \in T_0(Y);\ \|z\| \leq 1\}$. By Minkowski's Second Theorem we have

$$\lambda_1\cdots\lambda_{2\dim Y} \leq 2^{2\dim Y}\frac{\mathrm{vol}(\Omega_Y)}{\mu(2\dim Y)} \tag{3.1}$$

where $\mu(n) > 0$ denotes the Lebesgue volume of the unit ball in $\mathbb{R}^n$. There exist independent elements $\omega_1,\ldots,\omega_{2\dim Y} \in \Omega_Y$ with $\|\omega_i\| \leq \lambda_i \leq \lambda_{2\dim Y}$. Let $\rho = \rho(X,\mathcal{L}) > 0$ denote the minimal norm of a non-zero period in Ω_X. Using (3.1) and $\lambda_i \geq \rho$ we estimate

$$\|\omega_i\| \leq \lambda_{2\dim Y} \leq \frac{2^{2\dim Y}}{\mu(2\dim Y)\rho^{2\dim Y-1}}\mathrm{vol}(\Omega_Y).$$

The first inequality of (i) follows from Lemma 3.1 applied to Y. The second inequality follows easily from (3.1).

Now say $z = \omega' + y$ is as in part (ii) where $\omega' \in \Omega_X$ and $y \in T_0(Y)$. The periods $\omega_1,\ldots,\omega_{2\dim Y}$ generate $T_0(Y)$ as an $\mathbb{R}$-vector space. So $y = \alpha_1\omega_1 + \cdots + \alpha_{2\dim Y}\omega_{2\dim Y}$ for some $\alpha_1,\ldots,\alpha_{2\dim Y} \in \mathbb{R}$. For each i we fix $a_i \in \mathbb{Z}$ with $|\alpha_i - a_i| \leq 1/2$. Then $\omega'' = a_1\omega_1 + \cdots + a_{2\dim Y}\omega_{2\dim Y}$ is a period of X and

$$z = \omega' + \omega'' + \sum_{i=1}^{2\dim Y}(\alpha_i - a_i)\omega_i.$$

Part (ii) follows with $\omega = \omega' + \omega''$, the inequalities from (i), and the triangle inequality. □

Replacing $\mathcal{L}$ by another ample line bundle leads to a notion of degree that is comparable to the old one.

Lemma 3.3. *Let $\mathcal{M}$ be an ample line bundle on X. There exists a constant $c \geq 1$ depending only on $X, \mathcal{L}$, and $\mathcal{M}$ but not on Y such that $c^{-1} \deg_{\mathcal{L}} Y \leq \deg_{\mathcal{M}} Y \leq c \deg_{\mathcal{L}} Y$.*

Proof. To distinguish the norms and volumes coming from both line bundles we write $\|\cdot\|_{\mathcal{L}}$, $\|\cdot\|_{\mathcal{M}}$ and $\mathrm{vol}_{\mathcal{L}}$, $\mathrm{vol}_{\mathcal{M}}$. Let $\omega_1, \ldots, \omega_{2\dim Y}$ be as in Lemma 3.2(i), so $\|\omega_1\|_{\mathcal{L}} \cdots \|\omega_{2\dim Y}\|_{\mathcal{L}} \leq c \deg_{\mathcal{L}} Y$. As all norms on $T_0(X)$ are equivalent there exists $c' \geq 1$ with $\|v\|_{\mathcal{M}} \leq c'\|v\|_{\mathcal{L}}$ for all $v \in T_0(X)$. So $\|\omega_1\|_{\mathcal{M}} \cdots \|\omega_{2\dim Y}\|_{\mathcal{M}} \leq {c'}^{2\dim Y} \|\omega_1\|_{\mathcal{L}} \cdots \|\omega_{2\dim Y}\|_{\mathcal{L}}$ and Hadamard's inequality implies

$$\mathrm{vol}_{\mathcal{M}}(\Omega_Y) \leq \|\omega_1\|_{\mathcal{M}} \cdots \|\omega_{2\dim Y}\|_{\mathcal{M}} \leq c{c'}^{2\dim Y} \deg_{\mathcal{L}} Y.$$

The second inequality in the lemma follows from $\mathrm{vol}_{\mathcal{M}}(\Omega_Y) = (\deg_{\mathcal{M}} Y)/(\dim Y)!$. The first one follows by symmetry. □

Definition 3.4. *If A is a torsion coset of X which is the translate of an abelian subvariety Y of X by a torsion point, we define its* **arithmetic complexity** *as*

$$\Delta_{\mathrm{arith}}(A) = \min\{\textit{order of } T;\ A = T + Y \textit{ and } T \textit{ has finite order}\}$$

and its **complexity** *as*

$$\Delta(A) = \max\left\{\Delta_{\mathrm{arith}}(A), \deg_{\mathcal{L}} Y\right\} \geq 1$$

where $\deg_{\mathcal{L}} Y$ is the degree of Y with respect to $\mathcal{L}$.

We do not emphasise the choice of $\mathcal{L}$ in the complexity. According to Lemma 3.3 changing the line bundle leads to an arithmetic complexity which is comparable to the original one up to a controlled factor. For our application it is enough to fix once and for all a line bundle on the ambient abelian variety.

3.2. **The modular setting.** In this section we describe the special subvarieties of $Y(1)^n$ together with some additional definitions and notations that will be used in the sequel.

Let $j : \mathbb{H} \to Y(1)$ denote the j-function. By π we denote the cartesian power of this map

$$\pi : \mathbb{H}^n \to Y(1)^n.$$

Two-by-two real matrices with positive determinant act on $\mathbb{H}$ by fractional linear transformations. If $g \in \mathrm{GL}_2^+(\mathbb{Q})$ then the functions $j(z)$ and $j(gz)$ on $\mathbb{H}$ are related by a modular polynomial

$$\Phi_N(j(z), j(gz)) = 0$$

for a suitable positive integer $N = N(g)$ (in fact $N(g)$ is the determinant if g is scaled to have relatively prime integer entries; see [25, Ch. 5, §2]).

Definition 3.5. *A* **strongly special curve** *in $\mathbb{H}^n$ is the image of a map of the form*

$$\mathbb{H} \to \mathbb{H}^n, \quad z \mapsto (g_1 z, \ldots, g_n z)$$

where $g_1 = 1, g_2 \ldots, g_n \in \mathrm{GL}_2^+(\mathbb{Q})$.

By a **strict partition** we will mean a partition in which one designated part only is permitted to be empty.

Definition 3.6. *Let $R = (R_0, R_1, \dots, R_k)$ be a strict partition of $\{1, \dots, n\}$ in which R_0 is permitted to be empty (and $k = 0$ is permitted). For each index j we let $\mathbb{H}^{R_j}$ denote the corresponding cartesian product. A* **weakly special subvariety** *(of type R) of $\mathbb{H}^n$ is a product*

$$Y = \prod_{j=1}^{k} Y_j$$

where $Y_0 \in \mathbb{H}^{R_0}$ is a point and, for $j = 1, \dots, k$, Y_j is a strongly special curve in $\mathbb{H}^{R_j}$. We have $\dim Y = k$.

Definition 3.7. *A weakly special subvariety is called a* **special subvariety** *if each coordinate of Y_0 is a quadratic point of $\mathbb{H}$.*

With a quadratic $z \in \mathbb{H}$ we associate its **discriminant** $\Delta(z)$, namely $\Delta(z) = b^2 - 4ac$ where $aZ^2 + bZ + c$ is the minimal polynomial for z over $\mathbb{Z}$ with $a > 0$.

Definition 3.8. *The* **complexity** *of a special subvariety Y is defined to be*

$$\Delta(Y) = \max\big(\Delta(z), N(g)\big)$$

over all the coordinates z of Y_0 and all $g = g_k g_\ell^{-1} \in \mathrm{GL}_2^+(\mathbb{Q})$ where g_k, g_ℓ are involved in the definition of some constituent strongly special curve $Y_i, i \geq 1$.

Note that a weakly special subvariety has a certain number of "non-special conditions", namely the number of coordinates of Y_0 which are not quadratic, and is special just if this number is zero.

Further, weakly special subvarieties come in families. Given a strict partition $R = (R_0, R_1, \dots, R_k)$ we may form a new strict partition S in which the elements previously in R_0 are made into individual parts, the parts $R_1, \dots, R_k$ are retained, but S_0 is empty. Now a weakly special subvariety W of type R comes with a choice of some $\#R_j$ elements in $\mathrm{GL}_2^+(\mathbb{Q})$ for the parts R_j if $j \geq 1$. This same choice determines a weakly special subvariety T of type S which is in fact special (even **strongly special**, as there are no fixed coordinates). The variety W now lies in the family of weakly special subvarieties of T corresponding to choices for the fixed coordinates R_0. It is thus a family of weakly special subvarieties of T parameterised by $\mathbb{H}^{R_0}$. We will call the members of the family **translates** of the strongly special subvariety $T \subseteq \mathbb{H}^{R_1 \cup \dots \cup R_k}$ corresponding to the given elements in $\mathrm{GL}_2^+(\mathbb{Q})$, the space of translates being $\mathbb{H}^{R_0}$. The translate of T by $t \in \mathbb{H}^{R_0}$ we denote T_t.

We apply the same terminology to the images in $Y(1)^n$. In particular, we have the following.

Definition 3.9. *A* **weakly special subvariety** *of $Y(1)^n$ is the image $j(Y)$ where Y is a weakly special subvariety of $\mathbb{H}^n$, and is* **special** *if Y is special (for some or equivalently all possible choices for Y). The* **complexity** *of a special subvariety $T \subseteq Y(1)^n$, denoted $\Delta(T)$, is equal to the complexity of Y (any choice will give the same complexity due to the $\mathrm{SL}_2(\mathbb{Z})$ invariance).*

As observed the weakly special subvariety $Y \subseteq \mathbb{H}^n$ is a fibre in a family of weakly special subvarieties of some special subvariety T. Thus, the image under j of Y and the other translates are algebraic subvarieties of $j(T)$.

4. Geodesic-optimal subvarieties

Throughout this section and if not further specified let X be a mixed Shimura variety or a semi-abelian variety defined over $\mathbb{C}$.

The collection of weakly special subvarieties, like the collection of special subvarieties, is closed under taking irreducible components of intersection and contains the ambient variety, so there is a smallest weakly special subvariety $\langle A\rangle_{\mathrm{geo}}$ containing a subvariety $A \subseteq X$. We denote by

$$\delta_{\mathrm{geo}}(A) = \dim\langle A\rangle_{\mathrm{geo}} - \dim A$$

the **geodesic defect** of A.

Definition 4.1. *Let $V \subseteq X$ be a subvariety. A subvariety $A \subseteq V$ is said to be* **geodesic-optimal** *(for V in X) if there is no subvariety B with $A \subsetneq B \subseteq V$ such that*

$$\delta_{\mathrm{geo}}(B) \leq \delta_{\mathrm{geo}}(A).$$

As for the defect, the specification of V and X will generally be suppressed.

What we call "geodesic-optimal" has been termed "cd-maximal" (co-dimension maximal) in the multiplicative context by Poizat [39]; see also [4].

Again, if $A \subseteq V$ is geodesic-optimal then it is an irreducible component of $V \cap \langle A\rangle_{\mathrm{geo}}$. Further, as special subvarieties are weakly special, we have, for any V and $A \subseteq V$, $\langle A\rangle_{\mathrm{geo}} \subseteq \langle A\rangle$ and so $\delta_{\mathrm{geo}}(A) \leq \delta(A)$. In contrast to the defect, the geodesic defect of a singleton is always 0. Therefore, a singleton is geodesic-optimal for V if and only if it is not contained in a coset of positive dimension contained in V.

Definition 4.2. *We say that X has the* **defect condition** *if for any subvarieties $A \subseteq B \subseteq X$ we have*

$$\delta(B) - \delta_{\mathrm{geo}}(B) \leq \delta(A) - \delta_{\mathrm{geo}}(A).$$

Proposition 4.3. *The defect condition holds*

(i) *if $X = \mathbb{G}_m^n$ is an algebraic torus,*
(ii) *or if X is an abelian variety,*
(iii) *or if $X = Y(1)^n$.*

Proof. Let $A \subseteq B \subseteq X$ be as in Definition 4.2. For (i) let $B \subseteq \mathbb{G}_{\mathrm{m}}^n$ and let

$$L = \{(a_1,\ldots,a_n) \in \mathbb{Z}^n;\ x_1^{a_1}\cdots x_n^{a_n} \text{ is constant on } B\},$$
$$M = \{(a_1,\ldots,a_n) \in \mathbb{Z}^n;\ x_1^{a_1}\cdots x_n^{a_n} \text{ is constant and a root of unity on } B\}$$

be free abelian groups. Then $\operatorname{codim}\langle B\rangle = \operatorname{rank} M$ and $\operatorname{codim}\langle B\rangle_{\mathrm{geo}} = \operatorname{rank} L$, so that

$$\delta(B) - \delta_{\mathrm{geo}}(B) = \operatorname{rank} L/M$$

is the multiplicative rank of constant monomial functions on B. Such functions remain constant and multiplicatively independent on A.

To prove (ii) let A and B be nested subvarieties of X. The coset $\langle B\rangle_{\mathrm{geo}}$ is a translate of an abelian subvariety Y of X. Let us write $\varphi : X \to X/Y$ for the quotient morphism. We fix an auxiliary point $P \in A(\mathbb{C})$.

We remark that $\langle P\rangle + Y$ is a torsion coset that contains $P + Y$. As $P \in B(\mathbb{C})$ we also have $B \subseteq P + Y$ and thus $\langle B\rangle \subseteq \langle P\rangle + Y$. We apply φ, which has kernel Y, to find $\varphi(\langle B\rangle) \subseteq \varphi(\langle P\rangle) \subseteq \varphi(\langle A\rangle)$. We conclude

$$\dim\langle B\rangle - \dim\langle B\rangle_{\rm geo} = \dim\langle B\rangle - \dim Y = \dim\varphi(\langle B\rangle) \le \dim\varphi(\langle A\rangle) \tag{4.1}$$

where the fact that $\langle B\rangle$ contains a translate of Y and basic dimension theory are used for the second equality.

The torsion coset $\langle A\rangle$ is the translate of an abelian subvariety Z of X by a torsion point. The fibres of $\varphi|_{\langle A\rangle} : \langle A\rangle \to \varphi(\langle A\rangle)$ contain translates of $Y \cap Z$. Using dimension theory we find $\dim\varphi(\langle A\rangle) \le \dim\langle A\rangle - \dim Y \cap Z$. We observe $\dim Y \cap Z \ge \dim\langle A\rangle_{\rm geo}$ and so $\dim\varphi(\langle A\rangle) \le \dim\langle A\rangle - \dim\langle A\rangle_{\rm geo}$. Now let us combine this bound with (4.1) to deduce

$$\dim\langle B\rangle - \dim\langle B\rangle_{\rm geo} \le \dim\langle A\rangle - \dim\langle A\rangle_{\rm geo}.$$

This inequality enables us to conclude

$$\begin{aligned}\delta(B) &= \dim\langle B\rangle - \dim B\\ &\le \dim\langle A\rangle + \dim\langle B\rangle_{\rm geo} - \dim\langle A\rangle_{\rm geo} - \dim B\\ &= \delta_{\rm geo}(B) - \delta_{\rm geo}(A) + \delta(A),\end{aligned}$$

as desired.

In case (iii), $\delta(B) - \delta_{\rm geo}(B)$ is just the number of constant and non-special coordinates of $\langle B\rangle_{\rm geo}$. Then any weakly special subvariety containing A (which is non-empty) but contained in $\langle B\rangle_{\rm geo}$, and in particular $\langle A\rangle_{\rm geo}$, has also at least that many non-special constant coordinates. □

Conjecture 4.4. *Every mixed Shimura variety (and every weakly special subvariety) has the defect condition.*

Proposition 4.5. *Let X have the defect condition, e.g. X is an abelian variety or $Y(1)^n$, and let $V \subseteq X$ be a subvariety. An optimal subvariety for V is geodesic-optimal for V.*

Proof. Let $A \subseteq V$ be an optimal subvariety and consider a subvariety B with $A \subseteq B \subseteq V$ such that $\delta_{\rm geo}(B) \le \delta_{\rm geo}(A)$. Then $\delta(B) - \delta_{\rm geo}(B) \le \delta(A) - \delta_{\rm geo}(A)$ and so

$$\delta(B) = \delta_{\rm geo}(B) + \delta(B) - \delta_{\rm geo}(B) \le \delta_{\rm geo}(A) + \delta(A) - \delta_{\rm geo}(A) = \delta(A).$$

Since A is assumed optimal we must have $B = A$, and so A is geodesic optimal.

Finally, the proposition applies to abelian varieties and $Y(1)^n$ because of Proposition 4.3. □

5. Weak complex Ax

In this section we formulate various Ax-Schanuel type conjectures. In the context of abelian varieties, these conjectures will be theorems. But their modular counterparts are largely conjectural.

As a warming-up let us recall a consequence of "Ax-Schanuel" [2] in the complex setting.

Let now $A \neq \emptyset$ be an irreducible complex analytic subspace of some open $U \subseteq \mathbb{C}^n$ such that locally the coordinate functions $z_1, \ldots, z_n$ and $\exp(z_1), \ldots, \exp(z_n)$ are defined and meromorphic on A.

Definition 5.1. *The functions $z_1, \ldots, z_n$ on A are called* **linearly independent over $\mathbb{Q}$ mod $\mathbb{C}$** *if there are no non-trivial relations of the form $\sum_{i=1}^n q_i z_i = c$ on A where $q_i \in \mathbb{Q}$ and $c \in \mathbb{C}$.*

The linear independence of the z_i over $\mathbb{Q}$ mod $\mathbb{C}$ on A is equivalent to the multiplicative independence of the $\exp(z_i)$ on A.

Theorem 5.2 (Ax). *If the functions $z_1, \ldots, z_n$ on A are linearly independent over $\mathbb{Q}$ mod $\mathbb{C}$ then*

$$\operatorname{trdeg}_{\mathbb{C}} \mathbb{C}(z_1, \ldots, z_n, \exp(z_1), \ldots, \exp(z_n)) \geq n + \dim A.$$

5.1. **The abelian setting.** We will state two variants on Ax's Theorem that are sufficient to treat unlikely intersections in abelian varieties.

Let X be an abelian variety defined over $\mathbb{C}$. We write $T_0(X)$ for the tangent space of X at the origin. Moreover, there is a complex analytic group homomorphism $\exp : T_0(X) \to X(\mathbb{C})$. Here we use the symbol exp instead of π to emphasise the group structure.

Theorem 5.3 (Ax). *Let $U \subseteq T_0(X)$ be a complex vector subspace and $z \in T_0(X)$. Let K be an irreducible analytic subset of an open neighborhood of z in $z + U$. If B is the Zariski closure of $\exp(K)$ in X, then B is irreducible and*

$$\delta_{\text{geo}}(B) \leq \dim U - \dim K.$$

Proof. See Corollary 1 in [3]. □

The following statement is sometimes called the Ax-Lindemann-Weierstrass Theorem for abelian varieties. Theorem 5.3 will be used in Section 6.1 whereas Ax-Lindemann-Weierstrass makes its appearance near the end in Section 9 where we apply it in connection with a variant of the Pila-Wilkie Theorem. Ax-Lindemann-Weierstrass is sufficient to prove Theorem 1.1, but Theorem 1.3, situated in higher dimension, requires both statements. The reason seems to be that certain technical difficulties disappear in low dimension.

We may also consider $T_0(X)$ as a real vector space of dimension $2 \dim X$. After fixing an isomorphism $T_0(X) \cong \mathbb{R}^{2 \dim X}$ it makes sense to speak about semi-algebraic maps $[0, 1] \to T_0(X)$.

Theorem 5.4 (Ax). *Let $\beta : [0, 1] \to T_0(X)$ be real semi-algebraic and continuous with $\beta|_{(0,1)}$ real analytic. The Zariski closure in X of the image of $\exp \circ \beta$ is a coset.*

Proof. Clearly, we may assume that β is non-constant. The Zariski closure B of the image $\exp(\beta([0, 1]))$ is irreducible since $\exp \circ \beta$ is continuous and real analytic on $(0, 1)$.

By considering Taylor expansions around points of $(0, 1)$, the restriction $\beta|_{(0,1)}$ extends to a holomorphic map γ with target $T_0(X)$ and defined on a domain in $\mathbb{C}$ which contains $(0, 1)$. By analyticity the image of $\exp \circ \gamma$ lies in B and $\operatorname{trdeg}_{\mathbb{C}} \mathbb{C}(\exp \circ \gamma) \leq \dim B$.

As β is real algebraic on $[0, 1]$ we find $\operatorname{trdeg}_{\mathbb{C}} \mathbb{C}(\gamma) \leq 1$ and therefore,

$$\operatorname{trdeg}_{\mathbb{C}} \mathbb{C}(\gamma, \exp \circ \gamma) \leq \operatorname{trdeg}_{\mathbb{C}} \mathbb{C}(\gamma) + \operatorname{trdeg}_{\mathbb{C}} \mathbb{C}(\exp \circ \gamma) \leq 1 + \dim B.$$

Let us apply the one variable case of Ax's Theorem 3 [3] to the holomorphic function $t \mapsto \gamma(t + 1/2) - \gamma(1/2)$ defined in a neighborhood of $0 \in \mathbb{C}$. According to the inequality of transcendence degrees, the smallest abelian subvariety $H \subseteq X$

containing all values $\exp(\gamma(t+1/2)-\gamma(1/2))$ as t runs over U has dimension at most $\dim B$. Therefore, $B=\exp(\gamma(1/2))+H$ which is what we claimed. □

5.2. **A product of modular curves.** Now we suppose that $X=Y(1)^n$ and that $\pi:\mathbb{H}^n\to X(\mathbb{C})$ is the n-fold cartesian product of the j-function.

Let again $A\neq\emptyset$ be an irreducible complex analytic subspace of some open $U\subseteq\mathbb{H}^n$, so that locally the coordinate functions $z_1,\dots,z_n$ and $j(z_1),\dots,j(z_n)$ are defined and meromorphic on A, and we have a finite set $\{D_j\}$ of derivations with $\operatorname{rank}(D_j z_i)=\dim A$, the rank being over the field of meromorphic functions on A.

Definition 5.5. *The functions $z_1,\dots,z_n$ on A are called* **geodesically independent** *if no z_i is constant and there are no relations $z_i=gz_j$ where $i\neq j$ and $g\in\mathrm{GL}_2^+(\mathbb{Q})$.*

The geodesic independence of the z_i is equivalent to the $j(z_i)$ being "modular independent", i.e. non-constant and no relations $\Phi_N(j(z_k),j(z_\ell))$ where $k\neq\ell$ and $\Phi_N(X,Y)$ is a modular polynomial.

The following conjecture might be considered the analogue of "Ax-Schanuel" for the j-function in a complex setting.

Conjecture 5.6 (Complex "Modular Ax-Schanuel")**.** *In the above setting, if the z_i are geodesically independent then*

$$\operatorname{trdeg}_{\mathbb{C}}\mathbb{C}(z_1,\dots,z_n,j(z_1),\dots,j(z_n))\geq n+\dim A.$$

It evidently implies a weaker "two-sorted" version that, with the same hypotheses, we have the weaker conclusion

$$\operatorname{trdeg}_{\mathbb{C}}\mathbb{C}(z_1,\dots,z_n)+\operatorname{trdeg}_{\mathbb{C}}\mathbb{C}(j(z_1),\dots,j(z_n))\geq n+\dim A.$$

This conjecture is open beyond some special cases ("Ax-Lindemann" [32] and "Modular Ax-Logarithms" [21]).

We pursue now some more geometric formulations. To frame these we need some definitions.

Definition 5.7. *By an* **(algebraic) subvariety** *of $\mathbb{H}^n$ we mean an irreducible component (in the complex analytic sense) of $W\cap\mathbb{H}^n$ for some algebraic subvariety $W\subseteq\mathbb{C}^n$.*

Definition 5.8. *A subvariety $W\subseteq\mathbb{H}^n$ is called* **geodesic** *if it is defined by some number of equations of the forms*

$$z_i=c_i,\quad c_i\in\mathbb{C};\quad z_k=g_{k\ell}z_\ell,\quad g\in\mathrm{GL}_2^+(\mathbb{Q}).$$

These are the "weakly special subvarieties" in the Shimura sense.

Definition 5.9. *By a* **component** *we mean a complex-analytically irreducible component of $W\cap\pi^{-1}(V)$ where $W\subseteq\mathbb{H}^n$ and $V\subseteq X$ are algebraic subvarieties.*

Let A be a component of $W\cap\pi^{-1}(V)$. We can consider the coordinate functions z_i and their exponentials as elements of the field of meromorphic functions (at least locally) on A, and we can endow this field with $\dim A$ derivations in such a way that $\operatorname{rank}(Dz_i)=\dim A$. Then (with Zcl indicating the Zariski closure)

$$\dim W\geq\dim\operatorname{Zcl}(A)=\operatorname{trdeg}_{\mathbb{C}}\mathbb{C}(z_1,\dots,z_n),$$

$$\dim V\geq\dim\operatorname{Zcl}(\pi(A))=\operatorname{trdeg}_{\mathbb{C}}\mathbb{C}(j\circ z_1,\dots,j\circ z_n)$$

and the "two-sorted" Modular Ax-Schanuel conclusion becomes

$$\dim W + \dim V \geq \dim X + \dim A$$

provided that the functions z_i are geodesically independent.

This condition is equivalent to A being contained in a proper geodesic subvariety. Let us take U' to be the smallest geodesic subvariety containing A. Let $X' = \exp U'$, which is an algebraic subtorus of X, and put $W' = W \cap U'$, $V' = V \cap X'$. We can choose coordinates $z_i, i = 1, \ldots, \dim A$ which are linearly independent over $\mathbb{Q}$ mod $\mathbb{C}$ and derivations such that $\operatorname{rank}(Dz_i) = \dim A$. We get the following variant of Ax-Schanuel in this setting.

Conjecture 5.10 (Weak Complex Ax (WCA): Formulation A.)**.** *Let U' be a geodesic subvariety of U. Put $X' = \exp U'$ and let A be a component of $W \cap \pi^{-1}(V)$, where $W \subseteq U'$ and $V \subseteq X'$ are algebraic subvarieties. If A is not contained in any proper geodesic subvariety of U' then*

$$\dim A \leq \dim V + \dim W - \dim X'.$$

I.e. (and as observed still more generally by Ax [3]), the intersections of W and $\pi^{-1}(V)$ never have "atypically large" dimension, except when A is contained in a proper geodesic subvariety.

It is convenient to give a different (equivalent) formulation.

Definition 5.11. *Fix a subvariety $V \subseteq X$.*

- (i) *A* **component with respect to** *V is a component of $W \cap \pi^{-1}(V)$ for some algebraic subvariety $W \subseteq U$.*
- (ii) *If A is a component we define its* **defect** *by $\delta(A) = \dim \operatorname{Zcl}(A) - \dim A$.*
- (iii) *A component A with respect to V is called* **optimal** *(for V) if there is no strictly larger component B w.r.t. V with $\delta(B) \leq \delta(A)$.*
- (iv) *A component A with respect to V is called* **geodesic** *if it is a component of $W \cap \pi^{-1}(V)$ for some weakly special subvariety $W = \operatorname{Zcl}(A)$.*

Conjecture 5.12 (WCA: Formulation B.)**.** *Let $V \subseteq X$ be a subvariety. An optimal component for V is geodesic.*

Formulation B is the statement we need. However, the two formulations are equivalent, and the proof of their equivalence is purely formal and applies in the semiabelain setting and indeed quite generally.

Proof that formulation A implies formulation B. We assume Formulation A and suppose that the component A of $W \cap \pi^{-1}(V)$ is optimal, where $W = \operatorname{Zcl}(A)$. Suppose that U' is the smallest geodesic subvariety containing A, and let $X' = \pi(U')$. Then $W \subseteq U'$. Let $V' = V \cap X'$. Then A is optimal for V' in U', otherwise it would fail to be optimal for V in U. Since A is not contained in any proper geodesic subvariety of U' we must have

$$\dim A \leq \dim W + \dim V' - \dim X'.$$

Let B be the component of $\pi^{-1}(V')$ containing A. Then B is also not contained in any proper geodesic subvariety of U', so, by Formulation A,

$$\dim B \leq \dim V' + \dim \operatorname{Zcl}(B) - \dim X'.$$

But $\dim B = \dim V'$, whence $\dim \mathrm{Zcl}(B) = \dim X'$, and so $\mathrm{Zcl}(B) = X'$, and B is a geodesic component. Now

$$\delta(A) = \dim W - \dim A \geq \dim X' - \dim V' = \delta(B)$$

whence, by optimality, $A = B$. □

Proof that formulation B implies formulation A. We assume Formulation B. Let U' be a geodesic subvariety of U, put $X' = \pi(U)$. Suppose $V \subseteq X', W \subseteq U'$ are algebraic subvarieties and A is a component of $W' \cap \pi^{-1}(V')$ not contained in any proper geodesic subvariety of U'. There is some optimal component B containing A, and B is geodesic, but since A is not contained in any proper geodesic, B must be a component of $\pi^{-1}(V')$ with $\mathrm{Zcl}(B) = U'$ and we have

$$\dim W - \dim A \geq \delta(A) \geq \delta(B) = \dim X' - \dim V$$

which rearranges to what we want. □

As already remarked, WCA holds for (semi-)abelian varieties, by Ax [3] (see also [23]).

To conclude we note that a true "Modular Ax-Schanuel" should take into account the derivatives of j. A well-known theorem of Mahler [26] implies that j, j', j'' are algebraically independent over $\mathbb{C}$ as functions on $\mathbb{H}$, and it is well-known too that $j''' \in \mathbb{Q}(j, j', j'')$ (see e.g. [5]).

Conjecture 5.13 (Modular Ax-Schanuel with derivatives)**.** *In the setting of "Modular Ax-Schanuel" above, if z_i are geodesically independent then*

$$\mathrm{trdeg}_{\mathbb{C}}\mathbb{C}(z_i, j(z_i), j'(z_i), j''(z_i)) \geq 3n + \dim A.$$

6. A finiteness result for geodesic-optimal subvarieties

6.1. **The abelian case.** Suppose X is an abelian variety defined over $\mathbb{C}$. In this section we prove the following finiteness statement on geodesic-optimal subvarieties for a fixed subvariety of X. We recall that geodesic-optimal subvarieties were introduced in Definition 4.1.

Proposition 6.1. *For any subvariety $V \subseteq X$ there exists a finite set of abelian subvarieties of X with the following property. If A is a geodesic-optimal subvariety for V, then $\langle A\rangle_{\mathrm{geo}}$ is a translate of a member of the said set.*

Any positive dimensional, geodesic-optimal subvariety $A \subsetneq V$ is "μ-anormal maximal" for a certain μ in Rémond's terminology [44]. His Lemme 2.6 and Proposition 3.2 together imply that $\langle A\rangle_{\mathrm{geo}}$ is a translate of an abelian subvariety coming from a finite set that depends only on V. Thus our proof Proposition 6.1 can be regarded as an alternative approach to Rémond's Theorem using the language of o-minimal structures. The reader interested only in the proof of Theorem 1.1 can safely skip this section.

We retain the meaning of the symbol $T_0(X)$ from Section 3 and we further write $\exp : T_0(X) \to X(\mathbb{C})$ for the exponential map. It is a holomorphic group homomorphism between complex manifolds whose kernel is the period lattice of X. We choose a basis of the period lattice and identify $T_0(X)$ with $\mathbb{R}^{2g}$ as a real vector space. However, we continue to use both symbols $T_0(X)$ and $\mathbb{R}^{2g}$; the former is useful to emphasise the complex structure and the latter is required as an ambient set for an o-minimal structure.

The open, semi-algebraic set $(-1,1)^{2g}$ contains a fundamental domain for the period lattice $\mathbb{Z}^{2g} \subseteq \mathbb{R}^{2g}$ in real coordinates. Under the identification $\mathbb{R}^{2g} \cong T_0(X)$ we fixed above, we may consider $(-1,1)^{2g}$ as a domain in $T_0(X)$.

Let V be a subvariety of X. Then

$$\mathcal{V} = \exp|_{(-1,1)^{2g}}^{-1}(V(\mathbb{C}))$$

is a subset of $\mathbb{R}^{2g}$ and definable in $\mathbb{R}_{\mathrm{an}}$. But under the identification mentioned before, $\mathcal{V}$ is also a complex analytic subset of $(-1,1)^{2g} \subseteq T_0(X)$. Thus it is a complex analytic space. The interplay of these two points of view will have many consequences. For an in-depth comparision between complex and o-minimal geometry we refer to Peterzil and Starchenko's paper [30].

Indeed, suppose Z is an analytic subset of a domain in $T_0(X)$ and $z \in Z$. Then some open neighborhood of z in Z is definable in $\mathbb{R}_{\mathrm{an}}$ as Z is defined by the vanishing of certain holomorphic functions. So the dimension $\dim_z Z$ of Z at z as a set definable in $\mathbb{R}_{\mathrm{an}}$ is well-defined. But there is also the notion of the dimension of Z at z as a complex analytic space [19].

In this section we will add the subscript $\mathbb{C}$ to the dimension symbol to signify the dimension as a complex analytic space.

As $\dim \mathbb{C} = 2 = 2\dim_{\mathbb{C}} \mathbb{C}$ the following lemma not surprising.

Lemma 6.2. *Let Z be an analytic subset of a finite dimension $\mathbb{C}$-vector space. If $z \in Z$ then $\dim_z Z = 2\dim_{\mathbb{C},z} Z$.*

Proof. Locally at z the complex analytic space Z is a finite union of prime components, each of which is analytic in a neighborhood of z in Z. Without loss of generality we may assume that Z is irreducible at z. After shrinking Z further we may suppose that Z is irreducible, definable in $\mathbb{R}_{\mathrm{an}}$, and satisfies $\dim_z Z = \dim Z$. We fix a decomposition of Z into cells $D_1 \cup \cdots \cup D_N$ and write $Z' = Z \smallsetminus \bigcup_{\dim D_i < \dim Z} \overline{D_i}$ where the bar signifies closure in Z. Then Z' is an open and non-empty subset of Z. So it must contain a smooth point z' of the complex analytic space Z. Around z' we find $\dim_{z'} Z = 2\dim_{\mathbb{C},z'} Z = 2\dim_{\mathbb{C},z} Z$ since Z, as an analytic space, is equidimensional. But z' is contained in a cell D_i of dimension $\dim Z$. So $\dim Z = 2\dim_{\mathbb{C},z} Z$. Our claim now follows from $\dim Z = \dim_z Z$. □

We fix an isomorphism of $\mathrm{End}(T_0(X))$, the endomorphisms of $T_0(X)$ as a $\mathbb{C}$-vector space, with $\mathbb{R}^{2g^2}$ as an $\mathbb{R}$-vector space. Let $\mathcal{O} \subseteq \mathrm{End}(T_0(X))$ be definable in $\mathbb{R}_{\mathrm{an}}$ and satisfy the following 2 properties.

(i) We have $0 \in \mathcal{O}$.
(ii) If $M \in \mathcal{O}$ and if $Y \subseteq X$ is an abelian subvariety of X, then $T_0(Y) \cap \ker M$ is the kernel of an element in $\mathcal{O}$.

In particular, $\mathcal{O}$ contains an element whose kernel is the tangent space of any given abelian subvariety of X.

Suppose $0 \le r \le g$ is an integer. We set

$$F_r = \big\{(z,M) \in \mathcal{V} \times \mathcal{O};\ \dim_{\mathbb{C}} \ker M = r \text{ and for all } N \in \mathcal{O} \text{ with } \ker M \subsetneq \ker N : \\ \dim\ker M - \dim_z \mathcal{V} \cap (z + \ker M) < \dim\ker N - \dim_z \mathcal{V} \cap (z + \ker N)\big\}.$$

Then F_r is definable in $\mathbb{R}_{\mathrm{an}}$ by standard properties of o-minimal structures, for example by Proposition 1.5, chapter 4 [15] taking local dimensions is definable. We set

$$E_r = \big\{(z,M) \in F_r;\ \text{for all } M' \in \mathcal{O} \text{ with } \ker M' \subsetneq \ker M : \\ \dim_z \mathcal{V} \cap (z + \ker M') < \dim_z \mathcal{V} \cap (z+\ker M)\big\}$$

which is also definable in $\mathbb{R}_{\text{an}}$.

Ax's Theorem 5.3 for abelian varieties will be used in

Lemma 6.3. (i) *If $(z,M) \in E_r$ there is an abelian subvariety $Y \subseteq X$ with $T_0(Y) = \ker M$.*

(ii) *The set*

$$\{\ker M;\ (z,M) \in E_r\}$$

is finite.

(iii) *Let A be a geodesic-optimal subvariety for V and let $\langle A\rangle_{\text{geo}}$ be a translate of an abelian subvariety Y. If $r = \dim_{\mathbb{C}} Y$ and $M \in \mathcal{O}$ with $T_0(Y) = \ker M$ then $(z,M) \in E_r$ for some $z \in (-1,1)^{2g}$.*

Proof. Say $(z,M) \in E_r$ is as in (i). We apply Ax's Theorem 5.3 to $U = \ker M$. For this we fix a prime component (in the complex analytic sense) $K \subseteq \mathcal{V} \cap (z+U)$ with $\dim_{\mathbb{C},z} K = \dim_{\mathbb{C},z} \mathcal{V} \cap (z+U)$. By shrinking K to an open neighborhood of z we may assume that K is irreducible and definable in $\mathbb{R}_{\text{an}}$. Let $B \subseteq X$ be as in Ax's theorem.

If $\langle B\rangle_{\text{geo}}$ is a translate of the abelian subvariety $Y \subseteq X$ then $K \subseteq z + T_0(Y)$ and so

$$K \subseteq z + T_0(Y) \cap U.$$

Observe that $T_0(Y) \cap U$ is the kernel of an element in $\mathcal{O}$ by property (ii) above. By the definition of E_r we must have $T_0(Y) \cap U = U$. Therefore, $U \subseteq T_0(Y)$.

Next we prove that equality holds. Indeed, if $U \subsetneq T_0(Y)$, then we may test U against any $N \in \mathcal{O}$ with $\ker N = T_0(Y)$ in the definition of F_r. Therefore,

$$\begin{aligned} \dim U - \dim_z K &= \dim U - \dim_z \mathcal{V} \cap (z+U) \\ &< \dim T_0(Y) - \dim_z \mathcal{V} \cap (z + T_0(Y)) \\ &\le \dim T_0(Y) - \dim_z \exp|_{(-1,1)^{2g}}^{-1}(B(\mathbb{C})) \end{aligned}$$

where the final inequality required $B \subseteq V \cap (\exp(z) + Y)$. On passing to complex dimensions we obtain $\dim_{\mathbb{C}} U - \dim_{\mathbb{C},z} K < \dim_{\mathbb{C}} Y - \dim_{\mathbb{C},z} B = \delta_{\text{geo}}(B)$ which contradicts the conclusion of Ax's Theorem. So we must have $\ker M = U = T_0(Y)$ and part (i) follows.

Now we prove (ii) by showing that only finitely many possible kernels $\ker M$ can arise from $(z,M) \in E_r$. The image of E_r under the projection $T_0(X) \times \text{End}(T_0(X)) \to \text{End}(T_0(X))$ is definable in $\mathbb{R}_{\text{an}}$. By part (i) the kernel of M is the tangent space of an abelian subvariety of X. But X has at most countably many abelian subvarieties which leaves us with at most countably many possible kernels. We fix a $\mathbb{C}$-basis for $T_0(X)$ and identify each M with the corresponding $g \times g$ matrix. The Plücker coordinates of a submatrix of M with maximal rank are in a countable set of an appropriate projective space. Plücker coordinates are algebraic expressions in the entries of M. So we end up with an at most countable and definable subset on each member of some affine covering of projective space. But an at most countable and definable set is finite. So there are at most finitely many $\ker M$.

Let A and Y be as in (iii). Since A is a geodesic-optimal subvariety for V it must be an irreducible component of $V \cap \langle A \rangle_{\mathrm{geo}}$. Let us fix $z \in \mathcal{V}$ such that $\exp(z)$ is a smooth complex point of A that is not contained in any other irreducible component of $V \cap \langle A \rangle_{\mathrm{geo}}$.

We first prove $(z, M) \in F_r$ by contradiction. So suppose there exists $N \in \mathcal{O}$ with $T_0(Y) \subsetneq \ker N$ and

$$\dim T_0(Y) - \dim_z \mathcal{V} \cap (z + T_0(Y)) \geq \dim \ker N - \dim_z \mathcal{V} \cap (z + \ker N). \tag{6.1}$$

As in (i) we fix a prime component K of $\mathcal{V} \cap (z + \ker N)$ that passes through z with $\dim_{\mathbb{C},z} K = \dim_{\mathbb{C},z} \mathcal{V} \cap (z + \ker N)$ and is irreducible. Let B be the Zariski closure of $\exp(K)$, then Ax's Theorem implies $\delta_{\mathrm{geo}}(B) \leq \dim_{\mathbb{C}} \ker N - \dim_{\mathbb{C}} K$. As exp is locally biholomorphic our choice of z implies that z is a smooth point of the complex analytic set $\mathcal{V} \cap (z + T_0(Y))$ which has dimension $\dim_{\mathbb{C}} A$ at this point. So

$$\delta_{\mathrm{geo}}(A) = \dim_{\mathbb{C}} T_0(Y) - \dim_{\mathbb{C}} A \geq \dim_{\mathbb{C}} \ker N - \dim_{\mathbb{C}} K \geq \delta_{\mathrm{geo}}(B) \tag{6.2}$$

follows from (6.1) after dividing by 2.

By smoothness, the intersection $\mathcal{V} \cap (z + T_0(Y))$ has a unique prime component K' passing through z. The dimension inequality for intersections, cf. Chapter 5, §3 [19], implies

$$\dim_{\mathbb{C},z} K \cap (z + T_0(Y)) \geq \dim_{\mathbb{C},z} K + \dim_{\mathbb{C}} T_0(Y) - \dim_{\mathbb{C}} \ker N.$$

Inequality (6.1) and $\dim_{\mathbb{C},z} \mathcal{V} \cap (z + T_0(Y)) = \dim_{\mathbb{C}} A$ imply that the right-hand side is at least $\dim_{\mathbb{C}} A$. But $K \cap (z + T_0(Y)) \subseteq \mathcal{V} \cap (z + T_0(Y))$ and on comparing dimensions at z we find that $K \cap (z + T_0(Y))$, and a fortiori K, contains a neighborhood of z in K'. This implies $A \subseteq B$. But (6.2) and the fact that A is a geodesic-optimal subvariety forces $A = B$. So $\dim_{\mathbb{C}} K \leq \dim_{\mathbb{C}} A$ and (6.2) applied again yields $\dim_{\mathbb{C}} T_0(Y) \geq \dim_{\mathbb{C}} \ker N$, a contradiction.

Second, we will show $(z, M) \in E_r$. Suppose on the contrary that there is $M' \in \mathcal{O}$ with $\ker M' \subsetneq T_0(Y)$ and

$$\dim_z \mathcal{V} \cap (z + \ker M') \geq \dim_z \mathcal{V} \cap (z + T_0(Y)).$$

The set on the right is a complex analytic space, smooth at z, and contains the former. So $\mathcal{V} \cap (z + \ker M')$ and $\mathcal{V} \cap (z + T_0(Y))$ coincide on an open neighborhood of z in $(-1, 1)^{2g}$. Therefore, an open neighborhood of 0 in $A(\mathbb{C}) - \exp(z)$ is contained in the group $\exp(\ker M')$; here and below we use the Euclidean topology. Said group need not be algebraic or even closed, but it does contain an open, non-empty subset of the complex points of

$$\underbrace{(A - \exp(z)) + \cdots + (A - \exp(z))}_{\dim X \text{ terms}}.$$

This sum equals $\langle A - \exp(z) \rangle = Y = \exp(T_0(Y))$. Hence $T_0(Y) \subseteq \ker M'$, which is the desired contradiction. □

Proof of Proposition 6.1. We will work with $\mathcal{O} = \mathrm{End}(T_0(X))$. Suppose that A is a geodesic-optimal subvariety for V. Let us fix $M \in \mathcal{O}$ such that $\langle A \rangle_{\mathrm{geo}}$ is the translate of an abelian subvariety whose tangent space is $\ker M$. Then $\ker M$ lies in a finite set by Lemma 6.3 parts (ii) and (iii). So $\langle A \rangle_{\mathrm{geo}}$ is the translate of an abelian subvariety of X coming from a finite set. □

Recall that V is a subvariety of X. We brief discussion the connection between Proposition 6.1 and anomalous subvarieties as introduced by Bombieri, Masser, and Zannier [10] for subvarieties of the algebraic torus.

Definition 6.4. *With X and V as above, a subvariety $A \subseteq V$ is called* **anomalous** *if*

$$\dim A \geq \max\{1, \dim\langle A\rangle_{\text{geo}} + \dim V - \dim X + 1\}. \tag{6.3}$$

If in addition A is not contained in any strictly larger anomalous subvariety of V, then we call A **maximal anomalous**. *The complement in V of the union of all anomalous subvarieties of V is denoted by V^{oa}.*

Any maximal anomalous subvariety A of V is geodesic-optimal. Indeed, after enlargening there is a geodesic-optimal subvariety B for V with $A \subseteq B$ and $\delta_{\text{geo}}(B) \leq \delta_{\text{geo}}(A)$. So

$$\dim B \geq \dim\langle B\rangle_{\text{geo}} - \dim\langle A\rangle_{\text{geo}} + \dim A \geq \dim\langle B\rangle_{\text{geo}} + \dim V - \dim X + 1$$

due to (6.3). Since $\dim B \geq \dim A \geq 1$ we see that B is anomalous. As A is maximal anomalous we find $B = A$. So A is geodesic-optimal for V.

According to Proposition 6.1, $\langle A\rangle_{\text{geo}}$ is the translate of an abelian subvariety coming from a finite set which depends only on V. Let Y be such an abelian subvariety. By a basic result in dimension theory of algebraic varieties the set of points $V(\overline{K})$ at which $X \to X/Y$ restricted to V has a fibre greater or equal to some prescribed value is Zariski closed in V. It follows that V^{oa} is Zariski open in V; we may thus use the notation $V^{\text{oa}}(\overline{K})$ for $\overline{K}$-rational points in V^{oa}.

Openness of V^{oa} was previously known due to work of Rémond [44] and proved earlier in the toric setting by Bombieri, Masser, and Zannier [10].

We remark that V^{oa} is possibly empty. Moreover, $V^{\text{oa}} = \emptyset$ if and only if there exists a Y as above such that

$$\dim \varphi(V) < \min\{\dim X/Y, \dim V\}$$

where $\varphi : X \to X/Y$ is the canonical morphism.

6.2. **Mobius varieties.** Let $X = Y(1)^n$.

It is convenient to introduce a family subvarieties of $\mathbb{H}^n$ parameterised by choices of elements of $\mathbb{H}$ and $\mathrm{GL}_2^+(\mathbb{R})$ in which weakly special subvarieties are the fibres corresponding to the $\mathrm{GL}_2^+(\mathbb{R})$ parameters having their image under scaling in $\mathrm{SL}_2(\mathbb{R})$ lying in the image of $\mathrm{GL}_2^+(\mathbb{Q})$. Observe that this image is a countable set. In [32] these were termed "linear subvarieties" but the denotation "Mobius" seems to be more appropriate.

We take z_i as coordinates in $\mathbb{H}^n$ and $g_i, i = 2, \ldots, n$ as coordinates on $\mathrm{GL}_2^+(\mathbb{R})^{n-1}$ The family of Mobius curves in $\mathbb{H}^n$ is the locus

$$M^{\{1,\ldots,n\}} \subseteq \mathbb{H}^n \times \mathrm{GL}_2^+(\mathbb{R})^{n-1}$$

defined by the equations $z_i = g_i z_1, i = 2, \ldots, n$. We view this as a family of curves in $\mathbb{H}^n$ parameterised by $g = (g_2, \ldots, g_n) \in \mathrm{GL}_2^+(\mathbb{R})^{n-1}$. For a subset $R \subseteq \{1, \ldots, n\}$ we define M^R to be the family of Mobius curves on the product of factors of $\mathbb{H}^n$ over indices in R, parameterised by the corresponding factors of $\mathrm{GL}_2^+(\mathbb{R})^n$ excluding the smallest one which plays the role of z_1, which we will denote $\mathrm{GL}_2^+(\mathbb{R})^{R_i}$.

Let $R = (R_0, R_1, \ldots, R_k)$ be a strict partition as near Definition 3.6. Define the family of Mobius subvarieties of type R to be the locus

$$M^R \subseteq \mathbb{H}^n \times \mathbb{H}^{R_0} \times \prod_{i=1}^{k} \mathrm{GL}_2^+(\mathbb{R})^{R_i}$$

defined by equations placing the $\mathbb{H}^{R_i} \times \mathrm{GL}_2^+(\mathbb{R})^{R_i}$ point in M^{R_i} for $i = 1, \ldots, k$, and each R_0-coordinate in $\mathbb{H}^n$ is set equal to the corresponding coordinate in $\mathbb{H}^{R_0}$. For each choice of parameters

$$t \in M_R = \mathbb{H}^{R_0} \times \prod_{i=1}^{k} \mathrm{GL}_2^+(\mathbb{R})^{R_i}$$

the corresponding fibre M_t^R is a **Mobius subvariety** $\mathbb{H}^n$.

Like weakly special subvarieties, Mobius subvarieties come in families of "translates". For a fixed $g \in \prod_{i=1}^{k} \mathrm{GL}_2^+(\mathbb{R})^{R_i}$, the choices of $z \in \mathbb{H}^{R_0}$ give a family of Mobius subvarieties, the "translates" of the corresponding "strongly Mobius subvariety" M_g of the appropriate subproduct of $\mathbb{H}^n$, and the totality of the translates form a Mobius subvariety with no fixed coordinates.

A component $A \subseteq \mathbb{H}^n$ is contained in some smallest Mobius subvariety L_A, has a Mobius defect

$$\delta_M(A) = \dim L_A - \dim A$$

A component $A \subseteq j^{-1}(V)$ will be called **Mobius optimal** (for V in X) if there is no component B with $A \subsetneq B \subseteq j^{-1}(V)$ and $\delta_M(B) \leq \delta_M(A)$.

Proposition 6.5. *Assume WCA. Let $V \subseteq X$ be a subvariety. Then the set of*

$$g \in \prod_{i=1}^{k} \mathrm{GL}_2^+(\mathbb{R})^{R_i}$$

such that some translate of M_g intersects $j^{-1}(V)$ in a component which is Mobius optimal for V is finite modulo the action by $\prod_i \mathrm{SL}_2(\mathbb{Z})^{R_i}$.

Proof. The condition is $\prod_i \mathrm{SL}_2(\mathbb{Z})^{R_i}$ invariant, so the assertion is that such g come in finitely many $\prod_i \mathrm{SL}_2(\mathbb{Z})^{R_i}$ orbits. By WCA, any such g corresponds to a weakly special subvariety, and so the g in question belong to a countable set. However, every such g has a translate under $\prod_i \mathrm{SL}_2(\mathbb{Z})^{R_i}$ for which the optimal component has points of its full dimension in some fixed fundamental domain, say $\mathbb{F}_0^n$, and there the condition of optimality may be checked definably by considering dimensions of the intersection of $\pi^{-1}(V) \cap \mathbb{F}_0^n$ with Mobius subvarieties over the whole space of them, which is definable. Thus, there is a definable (in $\mathbb{R}_{\mathrm{an,exp}}$) countable and hence finite set of g which contains a representative of every $\prod_i \mathrm{SL}_2(\mathbb{Z})^{R_i}$ orbit of such g. □

As observed, the g above all correspond to weakly special families; however, every g corresponding to a weakly special family having a translate with a geodesic-optimal intersection will also appear in this set, as such g (by WCA) are in particular optimal among Mobius varieties. We conclude a modular version of Proposition 6.1.

Proposition 6.6. *Assume WMA. Let $V \subseteq X$ be a subvariety. Then there is a finite set of basic special subvarieties such that every weakly special subvariety* *which has a geodesic-optimal component in its intersection with V is a translate of one of these.*

Finally, we observe that special subvarieties of type R arise as fibres of M^R of points of M_R with suitable rationality properties. Specifically, the coordinates in $\mathrm{GL}_2^+(\mathbb{R})^{R_i}$ are rational and those in $\mathbb{H}^{R_0}$ are quadratic; let us call these "special points". Of course the same fibre arises from non-special choices of the parameter too.

Proposition 6.7. *There is an absolute constant $c > 0$ with the following property. Let $T \subseteq Y(1)^n$ be a special subvariety with complexity $\Delta(T)$ containing a point $P \in Y(1)^n$ with pre-image $Q \in \mathbb{F}_0^n$. Then there exists a "special point" $t \in M_R$ with*

$$H(t) \leq c\Delta(T)^{10}$$

such that M_t^R is a component of the pre-image of T and $Q \in M_t^R$.

Proof. This follows from Lemmas 5.2 and 5.3 of [21]. □

7. Counting semi-rational points

In this section we will work in a fixed o-minimal structure over $\mathbb{R}$. Our goal is to count points on a definable set where certain coordinates are algebraic of bounded height and degree and the rest are unrestricted. We will use our result to study unlikely intersections in abelian varieties.

Let us first fix some notation. Let $k \geq 1$ be an integer. We define the k-height of a real number $y \in \mathbb{R}$ as

$$H_k(y) = \min\big\{ \max\{|a_0|, \ldots, |a_k|\};\ a_0, \ldots, a_k \text{ are coprime integers, not all zero, with } a_0 y^k + \cdots + a_k = 0 \big\}$$

using the convention $\min \emptyset = +\infty$. A real number has finite k-height if and only if it has degree at most k over $\mathbb{Q}$. Let $m \geq 0$ be an integer. For $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$ we set

$$H_k(y) = \max\{H_k(y_1), \ldots, H_k(y_m)\}$$

and abbreviate $H(y) = H_1(y)$ if $y \in \mathbb{R}^m$. If $Z \subseteq \mathbb{R}^m$ is any subset, we define

$$Z(k, T) = \{y \in Z;\ H_k(y) \leq T\}$$

for $T \geq 1$.

If α is a mapping between two sets, then $\Gamma(\alpha)$ will denote the graph of α. Suppose $n \geq 0$ is an integer. A family parametrised by $\mathbb{R}^m$ is a subset $Z \subseteq \mathbb{R}^m \times \mathbb{R}^n$. In this case Z_y stands for the projection of $(\{y\} \times \mathbb{R}^n) \cap Z$ to $\mathbb{R}^n$ if $y \in \mathbb{R}^m$.

Let $Z \subseteq \mathbb{R}^m \times \mathbb{R}^n$ be a family parametrised by $\mathbb{R}^m$. Our goal is to determine the distribution of points $(y, z) \in Z$ where y has k-height at most T without restricting z. For a real number $T \geq 1$, we define

$$Z^{\sim}(k, T) = \{(y, z) \in Z;\ H_k(y) \leq T\}.$$

For technical reasons it is sometimes more convenient to work with

$$Z^{\sim,\mathrm{iso}}(k, T) = \{(y, z) \in Z^{\sim}(k, T);\ z \text{ is isolated in } Z_y\}.$$

Let $l \geq 0$ be an integer. We will use the second-named author's generalisation, stated below, of the Pila-Wilkie Theorem [35] to prove the following result for definable families. We refer to [32] for the definition of definable block and definably block family.

Theorem 7.1. *Let $F \subseteq \mathbb{R}^l \times \mathbb{R}^m \times \mathbb{R}^n$ be a definable family parametrised by $\mathbb{R}^l$ and $\epsilon > 0$. There is a finite number $J = J(F, k, \epsilon)$ of definable block families*

$$W^{(j)} \subseteq \mathbb{R}^{k_j} \times \mathbb{R}^l \times \mathbb{R}^m, \quad j \in \{1, \dots, J\}$$

parametrised by $\mathbb{R}^{k_j} \times \mathbb{R}^l$, for each such j a continuous, definable function

$$\alpha^{(j)} : W^{(j)} \to \mathbb{R}^n,$$

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

(i) *For all $j \in \{1, \dots, J\}$ and all $(t, x) \in \mathbb{R}^{k_j} \times \mathbb{R}^l$ we have*

$$\Gamma(\alpha^{(j)})_{(t,x)} \subseteq \{(y, z) \in F_x;\ z \text{ is isolated in } F_{(x,y)}\}.$$

(ii) *Say $x \in \mathbb{R}^l$ and $Z = F_x$. If $T \geq 1$ the set $Z^{\sim,\mathrm{iso}}(k, T)$ is contained in the union of at most cT^ϵ graphs $\Gamma(\alpha^{(j)})_{(t,x)}$ for suitable $j \in \{1, \dots, J\}$ and $t \in \mathbb{R}^{k_j}$.*

What follows is a useful corollary of the result above. Its assertion deals with $Z^\sim(k, T)$ and not $Z^{\sim,\mathrm{iso}}(k, T)$ which appears in the theorem.

Corollary 7.2. *Let F and ϵ be as in Theorem 7.1. We let π_1 and π_2 denote the projections $\mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$ and $\mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$, respectively. There exists a constant $c = c(F, k, \epsilon) > 0$ with the following property. Say $x \in \mathbb{R}^l$ and let $Z = F_x$. If $T \geq 1$ and $\Sigma \subseteq Z^\sim(k, T)$ with*

$$\#\pi_2(\Sigma) > cT^\epsilon,$$

there exists a continuous and definable function $\beta : [0, 1] \to Z$ such that the following properties hold.

(i) *The composition $\pi_1 \circ \beta : [0, 1] \to \mathbb{R}^m$ is semi-algebraic and its restriction to $(0, 1)$ is real analytic.*
(ii) *The composition $\pi_2 \circ \beta : [0, 1] \to \mathbb{R}^n$ is non-constant.*
(iii) *We have $\pi_2(\beta(0)) \in \pi_2(\Sigma)$.*
(iv) *If the o-minimal structure admits analytic cell decomposition, then $\beta|_{(0,1)}$ is real analytic.*

Here is the second-named author's counting theorem involving blocks.

Theorem 7.3 (Theorem 3.6 [32])**.** *Let $F \subseteq \mathbb{R}^l \times \mathbb{R}^m$ be a definable family parametrised by $\mathbb{R}^l$ and $\epsilon > 0$. There is a finite number $J = J(F, \epsilon)$ of definable block families*

$$W^{(j)} \subseteq \mathbb{R}^{k_j} \times \mathbb{R}^l \times \mathbb{R}^m, \quad j = 1, \dots, J,$$

each parametrised by $\mathbb{R}^{k_j} \times \mathbb{R}^l$, and a constant $c = c(F, k, \epsilon)$ with the following properties.

(i) *For all $(t, x) \in \mathbb{R}^{k_j} \times \mathbb{R}^l$ and all $j \in \{1, \dots, J\}$ we have $W_{(t,x)} \subseteq F_x$.*
(ii) *For all $x \in \mathbb{R}^l$ and $T \geq 1$ the set $F_x(k, T)$ is contained in the union of at most cT^ϵ definable blocks of the form $W^{(j)}_{(t,x)}$ for suitable $j = 1, \dots, J$ and $t \in \mathbb{R}^{k_j}$.*

Proof of Theorem 7.1. We refer to van den Dries's treatment of cells in Chapter 3 [15]. His convention for a cell $C \subseteq \mathbb{R}^m \times \mathbb{R}^n$ has the following advantage when considering it as a family parametrised by $\mathbb{R}^m$. If $y \in \mathbb{R}^m$ then $C_y \subseteq \mathbb{R}^n$ is either empty or a cell of dimension $\dim C - 1$.

We begin the proof of the theorem with a reduction step. Let us consider the set

$$F' = \{(x, y, z) \in F;\ z \text{ is isolated in } F_{(x,y)}\}.$$

We claim that it is definable. Indeed, let $C_1 \cup \cdots \cup C_N$ be a cell decomposition of F. Then $F_{(x,y)} = (C_1)_{(x,y)} \cup \cdots \cup (C_N)_{(x,y)}$ and each $(C_i)_{(x,y)}$ is either empty or a cell. The dimension of a non-empty $(C_i)_{(x,y)}$ does not depend on (x, y) and is the same locally at all points. Therefore, F' is a union of a subclass of the C_i and hence definable.

Since F' is definable it suffices to complete the proof with F replaced by F'. We thus assume that z is isolated in $F_{(x,y)}$ for all $(x, y, z) \in F$.

By general properties of an o-minimal structure, the number of connected components in a definable family is finite and bounded from above uniformly. So $\#F_{(x,y)} \leq c_1$ for all $(x, y) \in \mathbb{R}^l \times \mathbb{R}^m$ where c_1 is independent of x and y.

Let $\pi : \mathbb{R}^l \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^l \times \mathbb{R}^m$ denote the natural projection. Then $E^{(1)} = \pi(F)$ is a definable set. By Definable Choice, Chapter 6 Proposition 1.2(i) [15], there is a definable function $f^{(1)} : E^{(1)} \to \mathbb{R}^n$ whose graph $\Gamma(f^{(1)})$ lies in F. This graph is definable and so is $F \setminus \Gamma(f^{(1)})$.

The cardinality of a fibre of $F \setminus \Gamma(f^{(1)})$ considered as a family over $\mathbb{R}^l \times \mathbb{R}^m$ is at most $c_1 - 1$. If $F \neq \Gamma(f_1)$, Definable Choice yields a definable function $f^{(2)} : E^{(2)} \to \mathbb{R}^n$ on $E^{(2)} = \pi(F \setminus \Gamma(f_1))$ whose graph is inside $F \setminus \Gamma(f_1)$. The fibre above (x, y) of $F \setminus (\Gamma(f_1) \cup \Gamma(f_2))$ has at most $c_1 - 2$ elements.

This process exhausts all fibres of F after $c_2 \leq c_1$ steps. We get definable families $E^{(1)}, \ldots, E^{(c_2)} \subseteq \mathbb{R}^l \times \mathbb{R}^m$ parametrised by $\mathbb{R}^l$ and definable functions

$$f^{(i)} : E^{(i)} \to \mathbb{R}^n \quad \text{for} \quad i \in \{1, \ldots, c_2\} \quad \text{with} \quad \bigcup_{i=1}^{c_2} \Gamma(f^{(i)}) = F. \tag{7.1}$$

We can decompose each $E^{(i)}$ into finitely many cells on which $f^{(i)}$ is continuous. So after possibly increasing c_2 we may suppose that each $f^{(i)}$ is continuous and definable.

If $x \in \mathbb{R}^l$, not all coordinates of a point in $F_x^{\sim}(k, T)$ need to be algebraic. But the first m coordinates are and lead to points of k-height at most T on some $E_x^{(i)}$. These points can be treated using Theorem 7.3 applied to the family $E^{(i)}$. For every $i \in \{1, \ldots, c_2\}$ we obtain J_i definable block families $W^{(i,j)} \subseteq \mathbb{R}^{k_{i,j}} \times \mathbb{R}^l \times \mathbb{R}^m$ parametrised by $\mathbb{R}^{k_{i,j}} \times \mathbb{R}^l$ where $j \in \{1, \ldots, J_i\}$. They satisfy $W_{(t,x)}^{(i,j)} \subseteq E_x^{(i)}$ for $(t, x) \in \mathbb{R}^{k_{i,j}} \times \mathbb{R}^l$ and account for all points of k-height at most T on $E_x^{(i)}$.

Note that if $(t, x, y) \in W^{(i,j)}$, then $(x, y) \in E^{(i)}$. We consider the function

$$\alpha^{(i,j)} : W^{(i,j)} \to \mathbb{R}^n \quad \text{defined by} \quad (t, x, y) \mapsto f^{(i)}(x, y).$$

It is definable, being the composition of two definable functions: a projection and $f^{(i)}$. Moreover, $\alpha^{(i,j)}$ is continuous by our choice of the $E^{(i)}$. Observe $\Gamma(\alpha^{(i,j)})_{(t,x)} \subseteq F_x$ for all $(t, x) \in \mathbb{R}^{k_{i,j}} \times \mathbb{R}^l$ and this will yield (i).

Suppose $x \in \mathbb{R}^l$ and $(y, z) \in Z^{\sim}(k, T)$ with $Z = F_x$. The point $(x, y, z) \in F$ lies on the graph of some $f^{(i)}$ by (7.1). Hence $z = f^{(i)}(x, y)$ with $(x, y) \in E^{(i)}$. By definition we have $y \in E_x^{(i)}$. So $y \in E_x^{(i)}(k, T)$ since y has k-height at most T. Suppose $c(E^{(i)}, k, \epsilon)$ is the constant from Theorem 7.3. Then y is inside some $W_{(t,x)}^{(i,j)}$ where j and t are allowed to vary over $c(E^{(i)}, k, \epsilon)T^\epsilon$ possibilities. Therefore,

$(t,x,y) \in W^{(i,j)}$ and $(t,x,y,z) \in \Gamma(\alpha^{(i,j)})$ or equivalently, $(y,z) \in \Gamma(\alpha^{(i,j)})_{(t,x)}$. Part (ii) and the theorem follow after renumbering the $\alpha^{(i,j)}$ and $W^{(i,j)}$ using a single index. □

Proof of Corollary 7.2. The constant c from the this corollary comes from Theorem 7.1 applied to F, k, and ϵ.

Let $x \in \mathbb{R}^l$ and $Z = F_x \subseteq \mathbb{R}^m \times \mathbb{R}^n$. Suppose $T \geq 1$ satisfies

$$\#\pi_2(\Sigma) > cT^\epsilon \tag{7.2}$$

with $\Sigma \subseteq Z^\sim(k,T)$ as in the hypothesis.

First, let us assume $\Sigma \subseteq Z^{\sim,\mathrm{iso}}(k,T)$. By our Theorem 7.1 the set Σ is contained in the union of at most cT^ϵ graphs of continuous and definable functions. The Pigeonhole Principle and (7.2) yield two $(y,z),(y',z') \in \Sigma$ on the same graph with

$$z = \pi_2(y,z) \neq \pi_2(y',z') = z'. \tag{7.3}$$

From Theorem 7.3 we obtain a definable block family $W \subseteq \mathbb{R}^k \times \mathbb{R}^l \times \mathbb{R}^m$ and a continuous, definable function $\alpha : W \to \mathbb{R}^n$ with $(y,z),(y',z') \in \Gamma(\alpha)_{(t,x)}$ for a certain $t \in \mathbb{R}^k$. Moreover, $\Gamma(\alpha)_{(t,x)} \subseteq Z$.

The fibre $W_{(t,x)}$ is a definable block containing y and y'. A definable block is connected by definition. As above this means that there is a continuous, definable function $\gamma : [0,1] \to W_{(t,x)}$ with $\gamma(0) = y$ and $\gamma(1) = y'$. But $W_{(t,x)}$, being a definable block, is locally a semi-algebraic set. That is, for any $s \in [0,1]$ the point $\gamma(s)$ has a semi-algebraic neighborhood in $W_{(t,x)}$. Because $[0,1]$ is compact we may assume that γ is semi-algebraic.

We set

$$\beta(s) = (\gamma(s), \alpha(t,x,\gamma(s)))$$

and this yields a function $\beta : [0,1] \to Z$ which we show to satisfy all points in the assertion.

The function β is continuous and definable as γ and α possess these properties.

We note that $\pi_1 \circ \beta = \gamma$ is semi-algebraic by construction. This yields the first statement in (i).

We also note $\alpha(t,x,y) = z$, so $\beta(0) = (y,z) \in \Sigma$ and (iii) follows.

We find $\beta(1) = (y',z')$ in a analog manner. Therefore, (7.3) implies $\pi_2(\beta(0)) \neq \pi_2(\beta(1))$ and (ii) follows from this.

To complete the proof of (i) we use the fact that $\mathbb{R}_{\mathrm{alg}}$ admits analytic cell decomposition. There exist $0 = a_0 < a_1 < \cdots < a_{k+1} = 1$ such that each $\pi_1 \circ \beta|_{(a_i,a_{i+1})} : (a_i,a_{i+1}) \to \mathbb{R}^m \times \mathbb{R}^n$ is real analytic. By continuity and (ii) the restriction of $\pi_2 \circ \beta$ to some interval (a_i,a_{i+1}) is non-constant. If i is minimal with this property, then $\pi_2(\beta(a_i)) = \pi_2(\beta(0)) = z$. This will preserve (iii) after a linear reparametrisation of (a_i,a_{i+1}) to $(0,1)$. Thus we may suppose that $\pi_1 \circ \beta|_{(0,1)}$ is real analytic and this completes (i).

To prove (iv) we must assume that the ambient o-minimal structure admits analytic cell decomposition. As before we cover $[0,1]$ by finitely many open intervals and points, such that $\pi_2 \circ \beta$ restricted to each open interval is real analytic. We again linearly rescale the first open interval on which $\pi_2 \circ \beta$ is non-constant to $(0,1)$. So $\beta_{(0,1)} : (0,1) \to \mathbb{R}^m \times \mathbb{R}^n$ is real analytic.

Second and finally, let us suppose $\Sigma \not\subseteq Z^{\sim,\mathrm{iso}}(k,T)$. We fix any $(y,z) \in \Sigma \setminus Z^{\sim,\mathrm{iso}}(k,T)$; then $H_k(y) \leq T$ and the connected component of Z_y containing z has positive dimension. This component is definably connected. So we may fix a definable and continuous path $\alpha : [0,1] \to Z_y$ connecting $\alpha(0) = z$ with any other point $\alpha(1) \neq z$ of said component. Properties (i)-(iii) follow with the function $\beta(t) = (y, \alpha(t)) \in Z$ for $t \in [0,1]$. As above, rescaling implies (iv). □

Say $Z \subseteq \mathbb{R}^m$ is a definable. The corollary applied (with $m = n$) to the diagonal embedding $F = \{(z,z);\ z \in Z\} \subseteq \mathbb{R}^m \times \mathbb{R}^m$ recovers Pila and Wilkie's Theorem 1.8 [35].

8. Large Galois orbits

Let X be $Y(1)^n$ or an abelian variety defined over a field K which we take to be finitely generated over $\mathbb{Q}$. Recall that we have a notion of complexity of special subvarieties of X, cf. Sections 3.1 and 3.2 for the abelian and modular cases, respectively. Suppose $V \subseteq X$ is a subvariety defined over K. We consider various assertions.

(GO1) There exist $c, \eta > 0$ such that, if P is a point of V defined over a field extension of K, then

$$[K(P) : K] \geq c\Delta(\langle P \rangle)^\eta.$$

(GO2) There exist $c, \eta > 0$ such that if the singleton $\{P\} \subseteq V$ is an optimal subvariety for V, then

$$[K(P) : K] \geq c\Delta(\langle P \rangle)^\eta.$$

(GO3) There exist $c, \eta > 0$ such that, if W is an irreducible component of $V \cap \langle W \rangle$, then

$$[K(W) : K] \geq c\Delta(\langle W \rangle)^\eta.$$

Now GO3 is simply too strong: e.g. if $V = X = Y(1)^n$ then a component of $V \cap W$ is W and $\langle W \rangle = W$ for any special subvariety W. But some specials have large complexity and small Galois orbits. Indeed, any strongly special subvariety of $Y(1)^2$ is defined by a geometrically irreducible polynomial in integer coefficients. It would be interesting to know if GO3 is at least true in the case where W is a point.

Also GO1 seems to be very strong. Taking $V = X = Y(1)^n$ and $P \in \mathbb{Q}^n$, the existence of c, δ means that $\Delta(\langle P \rangle)$ is bounded as P runs over all rational points, so that only finitely many $\langle P \rangle$ arise. But this could be true. E.g. in $Y(1)^2$ we know that there are only finitely many modular curves with non-CM, non-cuspidal rational points by a result of Mazur. It is however stronger than we need.

For $V = Y(1)^n$ and if P is a special point, then statement of GO2 reverts to the conjecture on Galois orbits of special points. In this case GO2 follows from the Brauer-Siegel Theorem. However, we have no proof of the general case of GO2 at the moment. Also, GO1 is odd if P is transcendental, while in GO2 it must be algebraic over K; indeed, $\{P\}$ must be an irreducible component of $V \cap \langle P \rangle$ if P is an optimal singleton for V. For positive dimensional components in GO3 we must depend on o-minimality (and WCA) to bring us to finitely many families, and reduce to considering their translates.

Definition 8.1. *Let K be a field that is finitely generated over $\mathbb{Q}$. Suppose X is $Y(1)^n$ or an abelian variety defined over K and $V \subseteq X$ is a subvariety also defined over K. Let $s \geq 0$. We say that $LGO_s(V)$ is satisfied if there exists a constant*

$\kappa > 0$ with the following property. For any $P \in V(\overline{K})$ such that $\{P\}$ is an optimal singleton for V with $\dim\langle P\rangle \leq s$ we have

$$\Delta(\langle P\rangle) \leq (2[K(P):K])^{\kappa}. \tag{8.1}$$

If $r \geq 0$, we say that X satisfies LGO_s^r, if $LGO_s(V)$ is satisfied for all $V \subseteq X$ defined over $\overline{K}$ above with $\dim V \leq r$.

Finally, we say that X satisfies LGO if it satisfies LGO_s^r for all $r, s \geq 0$.

Conjecture 8.2. *Let K be finitely generated over $\mathbb{Q}$. If X is an abelian variety defined over K or if $X = Y(1)^n$, then X satisfies LGO.*

One might expect an analog conjecture to hold in any mixed Shimura variety (or perhaps even any weakly special subvariety thereof).

Suppose K is a number field. For special points of Shimura varieties the best results known are those of Tsimerman [46]: lower bounds of the above form for the size of the Galois orbit for special points in the coarse moduli space of principally polarised abelian varieties of dimension g, $\mathcal{A}_g$, for $g \leq 6$ (or on GRH for all g; see also [48]). For unlikely intersections of curves with special subvarieties of $Y(1)^n$ partial results are obtained in [21].

To prove a uniform version (which one could frame as a ZP for $Y(1)^n \times \mathbb{C}^k$ where the second factor is viewed with rational structure, as done in [33]), one would want that, for a family of subvarieties $\{V_t\}$, i.e. the fibres of some $V \subseteq Y(1)^n \times \mathbb{C}^k$, the constant κ for V_t in the family depends only on $[\mathbb{Q}(V_t):\mathbb{Q}]$, i.e. on the degree of the parameter t over $\mathbb{Q}$.

9. Unlikely intersections in abelian varieties

9.1. **The arithmetic complexity of a torsion coset.** In this section we will prove an upper bound on the arithmetic complexity of a torsion coset as introduced in Definition 3.4. Our main tool is a height lower bound due to Masser.

Suppose X is an abelian variety of dimension $g \geq 1$ defined over a number field K and $\mathcal{L}$ is an ample line bundle on X. To simplify notation we will suppose that K is a subfield of $\mathbb{C}$. For example, this enables us to consider the tangent space $T_0(X)$ as a $\mathbb{C}$-vector space. We let $\overline{K}$ denote the algebraic closure of K in $\mathbb{C}$.

After replacing $\mathcal{L}$ by $\mathcal{L} \otimes [-1]^*\mathcal{L}$, where $[-1]$ denotes inversion on the algebraic group X, we may assume that $\mathcal{L}$ is symmetric. Let $\hat{h}$ be the Néron-Tate height on $X(\overline{K})$ attached to $\mathcal{L}$. We recall that the group of homomorphisms between two abelian varieties is a finitely generated, free abelian group. Let d be the dimension of an abelian subvariety of X. For the next proposition we require

$$\lambda_X(d) = \sup\{\operatorname{rank}\operatorname{Hom}(X, X/H)\cdot\dim(X/H);\ H \subseteq X \text{ is an abelian subvariety over } \overline{K} \text{ with } \dim H = d\} < +\infty$$

where rank denotes the rank of a free abelian group and $\operatorname{Hom}(\cdot,\cdot)$ the group of homomorphisms over $\overline{K}$.

We observe that $\lambda_X(d) = 0$ if and only if $d = \dim X$.

Proposition 9.1. *There exists a constant $c > 0$ depending on X and $\mathcal{L}$ such that*

$$\Delta_{\text{arith}}(\langle P\rangle) \leq c[K(P):K]^{6g+1}$$

and

$$\deg_{\mathcal{L}} H \leq c[K(P):K]^{60g^4}\max\{1, \hat{h}(P)\}^{\lambda_X(\dim\langle P\rangle)}$$

for all $P \in X(\overline{K})$ where $H = \langle P\rangle - P$. In particular,

$$\Delta(\langle P\rangle) \le c[K(P):K]^{60g^4}\max\{1,\hat{h}(P)\}^{\lambda_X(\dim\langle P\rangle)}.$$

Bombieri, Masser, and Zannier [9] employed essentially best-possible height lower bound due to Amoroso and David [1] to prove a weak form of the Zilber-Pink Conjecture for curves in $\mathbb{G}_m^n$. Later, Rémond [42] developed this approach for abelian varieties using the geometry of numbers. We will follow a similar line of thought in this section. However, lower bounds of the same quality as Amoroso and David's result are not known for a general abelian variety. One advantage of the o-minimal approach is that it can cope with a height lower bound as long as it is polynomial in the reciprocal of the degree. Below, we will use such a lower bound due to Masser [27] that applies to any abelian variety defined over a number field. For the sake of simplicity we state the estimates in a form that is weaker than what Masser proved.

Theorem 9.2 (Masser). *There exists a constant $c > 0$ depending only on X, K, and $\mathcal{L}$ with the following property. Suppose $P \in X(\overline{K})$ and $D = [K(P):K]$ with $\hat{h}(P) < c^{-1}D^{-2g-9}$, then P is a torsion point of order at most cD^{6g+1}.*

Proof. This follows from the main theorem of [27] and the comments that follow it. □

Given $P \in X(\overline{K})$, a first consequence of Masser's theorem is a lower bound for the $[K(P):K]$ in terms of the arithmetic complexity of $\langle P\rangle$. We do not yet need the Néron-Tate height.

Lemma 9.3. *There exists a constant $c_3 > 0$ such that $\Delta_{\mathrm{arith}}(\langle P\rangle) \le c_3[K(P):K]^{6g+1}$ for all $P \in X(\overline{K})$.*

Proof. We note that the conclusion becomes stronger when replacing K by a field extension. So we may assume that all abelian subvarieties of X are defined over K.

Bertrand proved that there is an integer $c_1 \ge 1$ such that any abelian subvariety $Y \subseteq X$ has a companion abelian subvariety $Z \subseteq X$ with $Y + Z = X$ such that $Y \cap Z$ is finite and contains at most c_1 elements. Ratazzi and Ullmo published a proof [41] of Bertrand's Theorem. The point here is that c_1 does not depend on Y.

Suppose that $\langle P\rangle$ is a translate of Y by a torsion point. Let us write $P = Q + R$ with $Q \in Y(\overline{K})$ and $R \in Z(\overline{K})$. A positive multiple of P lies in $Y(\overline{K})$. This and $\#Y \cap Z < \infty$ imply that R has finite order, say M. Masser's theorem implies $M \le c_2[K(R):K]^{6g+1}$.

By definition we have $\Delta_{\mathrm{arith}}(\langle P\rangle) \le M$ and thus $\Delta_{\mathrm{arith}}(\langle P\rangle) \le c_2[K(R):K]^{6g+1}$. It remains to bound $[K(R):K]$ from above in terms of $[K(P):K]$.

Suppose $\sigma, \sigma' \in \mathrm{Gal}(\overline{K}/K)$ with $\sigma(R) \ne \sigma'(R)$. If $\sigma(P) = \sigma'(P)$ then $P = Q+R$ yields $\sigma'(Q) - \sigma(Q) = \sigma(R) - \sigma'(R)$. This point is in $Y \cap Z$ as Y and Z are defined over K. This leaves at most c_1 possibilities for $\sigma(R) - \sigma'(R)$. We conclude

$$\begin{aligned}[K(P):K] &= \#\{\sigma(P);\ \sigma \in \mathrm{Gal}(\overline{K}/K)\}\\ &\ge \frac{1}{c_1}\#\{\sigma(R);\ \sigma \in \mathrm{Gal}(\overline{K}/K)\}\\ &= \frac{[K(R):K]}{c_1}.\end{aligned}$$

The lemma follows from the lower bound for $[K(R):K]$ we obtain above. □

We setup some additional notation before we come to the proof of Proposition 9.1.

Say Y is a second abelian variety, also defined over K, equipped with an ample and symmetric line bundle. Thus we have another Néron-Tate height $\hat{h}_Y : Y(\overline{K}) \to [0,\infty)$. We will assume that all elements in $\mathrm{Hom}(X,Y)$ are already defined over K. We set $\rho = \mathrm{rank}\,\mathrm{Hom}(X,Y)$. To avoid trivialities we shall assume $\rho \geq 1$, so in particular $\dim Y \geq 1$. We set

$$\mathrm{Hom}(X,Y)^*_{\mathbb{R}} = \{\psi \in \mathrm{Hom}(X,Y)\otimes\mathbb{R};\ \text{there is } \varphi \in \mathrm{Hom}(Y,X)\otimes\mathbb{R} \text{ with } \psi\varphi = 1\}$$

and fix a norm $\|\cdot\|$ on the finite dimensional vector space $\mathrm{Hom}(X,Y)\otimes\mathbb{R}$. For example, we could take the norm induced by the Rosati involution coming from the two line bundles. We consider $\mathrm{Hom}(X,Y)\otimes\mathbb{R}$ with the topology induced by $\|\cdot\|$. An element $\psi \in \mathrm{Hom}(X,Y)\otimes\mathbb{R}$ lies in $\mathrm{Hom}(X,Y)^*_{\mathbb{R}}$ precisely when the linear map $\mathrm{Hom}(Y,X)\otimes\mathbb{R} \to \mathrm{Hom}(Y,Y)\otimes\mathbb{R}$ given $\varphi \mapsto \psi\varphi$ is surjective. Therefore, $\mathrm{Hom}(X,Y)^*_{\mathbb{R}}$ is an open, possibly empty, subset of $\mathrm{Hom}(X,Y)\otimes\mathbb{R}$.

Below, $c_4, c_5, \ldots$ denote positive constants that depend only on these two abelian varieties, K, and the choosen line bundles.

The upper bound for the geometric part of the complexity involves the Néron-Tate height.

Lemma 9.4. *Suppose $Q > 1$ and let $P \in X(\overline{K})$ be in the kernel of a surjective element of* $\mathrm{Hom}(X,Y)$. *There is a surjective $\varphi \in \mathrm{Hom}(X,Y)$ with*

$$\hat{h}_Y(\varphi(P)) \leq c_4 Q^{-2/\rho}\hat{h}(P) \quad \textit{and} \quad \|\varphi\| \leq c_4 Q.$$

Proof. By Lemmas 2 and 5 [20] there is a compact subspace $\mathcal{K} \subseteq \mathrm{Hom}(X,Y)^*_{\mathbb{R}}, \varphi \in \mathrm{Hom}(X,Y), \varphi_0 \in \mathcal{K}$, and an integer q with $1 \leq q \leq Q$ such that $\hat{h}_Y(\varphi(P)) \leq c_4 Q^{-2/\rho}\hat{h}(P)$ and

$$\|q\varphi_0 - \varphi\| \leq c_4 Q^{-1/\rho}. \tag{9.1}$$

We emphasise that c_4 does not depend on P or Q.

The norm is bounded from above on the compact space $\mathcal{K}$. So we obtain $\|\varphi\| \leq \|q\varphi_0 - \varphi\| + q\|\varphi_0\| \leq c_4 Q$ after increasing c_4, if necessary.

We must show that φ is surjective. Recall that $\mathrm{Hom}(X,Y)^*_{\mathbb{R}}$ is open. Therefore, there is $Q_0 \geq 1$ such that $Q > Q_0$ implies $\varphi/q \in \mathrm{Hom}(X,Y)^*_{\mathbb{R}}$ by (9.1). In particular φ has a right inverse in $\mathrm{Hom}(Y,X)\otimes\mathbb{R}$. In this case it must already have a right inverse in $\mathrm{Hom}(Y,X)\otimes\mathbb{Q}$ by basic linear algebra. So φ is surjective.

Now if $Q \leq Q_0$ it suffices to fixed in advance a surjective homomorphism for φ. By general properties of the Néron-Tate height $\hat{h}_Y(\varphi(P))$ is bounded from above linearly in $\hat{h}(P)$, cf. expression (8) in Section 2 [20]. The lemma follows after increasing c_4 a final time. □

Lemma 9.5. *Suppose $P \in X(\overline{K})$ is in the kernel of a surjective element of* $\mathrm{Hom}(X,Y)$ *and $D = [K(P) : K]$. There exists a surjective $\varphi \in \mathrm{Hom}(X,Y)$ with $\varphi(P) = 0$ and*

$$\|\varphi\| \leq c_6 D^{6\dim Y + 1 + (2\dim Y + 9)\rho/2}\max\{1, \hat{h}(P)\}^{\rho/2}.$$

Proof. Suppose P is as in the hypothesis and let us abbreviate $h = \max\{1, \hat{h}(P)\}$. We suppose that φ is a surjective morphism with $\varphi(P) = 0$ and with $\|\varphi\|$ minimal among all such morphisms. Let $c_5 > 0$ be the constant from Masser's Theorem applied to Y. We will assume

$$\|\varphi\| > 2(c_4c_5)^{1+\rho/2}D^{6\dim Y+1+(2\dim Y+9)\rho/2}h^{\rho/2}, \tag{9.2}$$

with c_4 from Lemma 9.4, and derive a contradiction. This implies the proposition with $c_6 = 2(c_4c_5)^{1+\rho/2}$.

Let us define the integer

$$N = [c_5 D^{6\dim Y+1}].$$

Without loss of generality we have $c_4 \geq 1$ and $c_5 \geq 1$, so $N \geq 1$. We define further

$$Q = \frac{\|\varphi\|}{2c_4N}.$$

Our assumption implies $\|\varphi\| > 2c_4c_5D^{6\dim Y+1} \geq 2c_4N$ and so $Q > 1$. We apply Lemma 9.4 to find a surjective $\phi \in \mathrm{Hom}(X, Y)$ with $\hat{h}_Y(\phi(P)) \leq c_4Q^{-2/\rho}h$ and $\|\phi\| \leq c_4Q$.

Say $1 \leq n \leq N$, then

$$\|n\phi\| \leq N\|\phi\| \leq c_4NQ = \frac{\|\varphi\|}{2} < \|\varphi\|.$$

By minimality of $\|\varphi\|$ we conclude $n\phi(P) \neq 0$. So $\phi(P)$ is either non-torsion or has finite order strictly greater than N. Masser's Theorem excludes the second alternative and provides $\hat{h}_Y(\phi(P)) \geq c_5^{-1}D^{-2\dim Y-9}$. We combine this bound with the upper bound from Lemma 9.4 to deduce $c_4c_5D^{2\dim Y+9}h \geq Q^{2/\rho}$. Inserting our choice for Q and N gives

$$c_4c_5D^{2\dim Y+9}h \geq \left(\frac{\|\varphi\|}{2c_4N}\right)^{2/\rho} \geq \left(\frac{\|\varphi\|}{2c_4c_5D^{6\dim Y+1}}\right)^{2/\rho}.$$

The incompatibility with (9.2) is the desired contradiction. □

Proof of Proposition 9.1. The bound for the arithmetic part of the complexity follows from Lemma 9.3. The complexity of $\langle P\rangle$ is the maximum of $\Delta_{\mathrm{arith}}(\langle P\rangle)$ and $\deg_{\mathcal{L}} Y$ and so it suffices to prove the second bound.

Without loss of generality we may assume $H \neq X$. By Poincaré's Complete Reducibility Theorem 5.3.5 [6] there are up-to $\overline{K}$-isogeny only finitely many possibilities for X/H. So we may assume that there is an abelian variety Y, coming from a finite set independent of P, and a surjective homomorphism $X \to Y$ whose kernel contains H as a connected component. After multiplying said homomorphism by a positive integer we may assume P lies in its kernel. We observe that the assertion of the proposition becomes stronger when enlarging K, so we may assume that H, X, and all elements in $\mathrm{Hom}(X, Y)$ are defined over K.

We apply Lemma 9.5 to find a surjective homomorphism $\varphi : X \to Y$ with $\varphi(P) = 0$ and whose norm is bounded from above by $c_6D^{6\dim Y+1+(2\dim Y+9)\rho/2}h^{\rho/2}$ with $D = [K(P) : K]$ and $\rho > 0$ the rank of $\mathrm{Hom}(X, Y)$. We have $\dim H = \dim X - \dim Y$ by a dimension counting argument.

Let $\Omega_X \subseteq T_0(X)$ denote the period lattice and tangent space of X at the origin. We use the same norm $\|\cdot\|$ on $T_0(X)$ as introduced in Section 3.1. If $\Omega_Y \subseteq T_0(Y)$ denotes the period lattice of Y, then φ induces a linear map $\Omega_X \to \Omega_Y$. Say $g' = \dim H$.

To proceed we apply an adequate version of Siegel's Lemma to solve $\varphi(\omega_i) = 0$ in linearly independent periods $\omega_1, \ldots, \omega_{2g'} \in \Omega_X$ with controlled norm. Indeed, we may refer to Corollary 2.9.9 [7], however the numerical constants there will not matter for us. Siegel's Lemma yields the first inequality in

(9.3)
$$\|\omega_1\| \cdots \|\omega_{2g'}\| \leq c_7 \|\varphi\|^{2(g-g')} = c_7 \|\varphi\|^{2 \dim Y} \leq c_8 D^{58g^4} \max\{1, \hat{h}(P)\}^{\rho \dim Y},$$

the second one follows from the bound for $\|\varphi\|$ we deduced further up and

$$(12 \dim Y + 2 + (2 \dim Y + 9)\rho) \dim Y \leq (12g + 2 + (2g + 9)4g^2)g \leq 58g^4$$

as $\rho \leq 4g \dim Y$ and $\dim Y \leq g$.

Lemma 3.1 and Hadamard's Inequality yield

$$[\ker \varphi : H] \deg_{\mathcal{L}} H = (\dim H)! [\ker \varphi : H] \mathrm{vol}(\Omega_H) \leq (\dim H)! \|\omega_1\| \cdots \|\omega_{2g'}\|.$$

As $[\ker \varphi : H] \geq 1$ we get an upper bound for $\deg_{\mathcal{L}} H$ which yields the assertion when combined with (9.3). □

9.2. **LGO and the Néron-Tate height.** We begin by exhibiting a connection between LGO, Definition 8.1, and height upper bounds on abelian varieties

Let X be an abelian variety defined over a number field K and suppose $\mathcal{L}$ is a symmetric, ample line bundle on X. Let $\hat{h}$ denote the associated Néron-Tate height function. Observe that any optimal subvariety for a subvariety of X defined over K is defined over $\overline{K}$.

Definition 9.6. *Let $V \subseteq X$ be a subvariety defined over K. Let $S \geq 0$. We define $\mathrm{Opt}(V; S)$ to be the set of those $A \in \mathrm{Opt}(V)$ which contain a $P \in A(\overline{K})$ with $\hat{h}(P) \leq (2[K(A) : K])^S$; here $K(A)$ denotes the smallest subfield of $\overline{K}$ containing K over which A is defined.*

In order to apply the counting strategy to study optimal subvarieties we must find a polynomial upper bound for the complexity of a torsion coset in terms of its arithmetic degree. Since the inequality in Proposition 9.1 also involves the height we make the following observation.

Proposition 9.7. *Let $V \subseteq X$ be a subvariety defined over K and let $s \geq 0$. Then $LGO_s(V)$ is satisfied if there exist $\epsilon > 0$ and $S \geq 0$ such that*

$$\hat{h}(P) \leq (2[K(P) : K])^S \left(\deg_{\mathcal{L}}(\langle P \rangle - P)\right)^{\frac{1}{\lambda_X(\dim\langle P \rangle)} - \epsilon}$$

for all optimal singletons $\{P\} \subseteq V$ with $\lambda_X(\dim\langle P \rangle) > 0$ and $\dim\langle P \rangle \leq s$. In particular, $LGO_s(V)$ is satisfied if the Néron-Tate height of an optimal singleton for V with defect at most s is bounded from above uniformly.

Proof. Observe that $\lambda_X(\dim\langle P \rangle) = 0$ if and only if $\langle P \rangle = X$. In this case $\{P\}$ can only be an optimal singleton for V if $V = \{P\}$ in which case the claim is trivial.

If $\lambda_X(\dim\langle P \rangle) > 0$ the claim is direct consequence of Proposition 9.1. □

The main result of this section states that a lower bound for the Galois orbit as in (8.1) is sufficient to prove that there are only finitely many optimal subvarieties for V. Although we believe (8.1) to always hold, we are not able to prove it. However, we can show unconditionally that $\mathrm{Opt}(V; S)$ is finite for any fixed $S \geq 0$.

Theorem 9.8. *Let X be an abelian variety defined over a field K which is finitely generated over $\mathbb{Q}$. Let $V \subseteq X$ be a subvariety defined over K.*

(i) *Say $r, s \geq 0$ and suppose that all quotients of X defined over a finite extension of K satisfy LGO^r_s. Then*

$$\{A \in \mathrm{Opt}(V);\ \mathrm{codim}_V A \leq r \text{ and } \dim\langle A\rangle - \dim\langle A\rangle_{\mathrm{geo}} \leq s\} \tag{9.4}$$

is finite.

(ii) *If K is a number field, then $\mathrm{Opt}(V; S)$ is finite for all $S \geq 0$.*

We obtain 2 corollaries, the first one follows from Lemma 2.7 which states that Conjectures 2.2 and 2.6 are equivalent.

Corollary 9.9. *Let us suppose that the height bound in Proposition 9.7 holds for all subvarieties of all abelian varieties defined over any number field. Then the Zilber-Pink Conjecture 2.2 holds for all subvarieties of all abelian varieties defined over any number field.*

Corollary 9.10. *Let X and V be as in Theorem 9.8. We suppose that all quotients of X defined over a finite extension of K satisfy LGO^r_s for all $r, s \geq 0$. Then $\mathrm{Opt}(V)$ is finite for any subvariety V of X defined over $\overline{K}$.*

Let us look more closely at the case when K is a number field and s is small. The Néron-Tate height of a torsion point vanishes. So by Proposition 9.7 any abelian variety over a number field satisfies LGO^r_0 for all $r \geq 0$. Part (i) of the theorem implies that V contains only finitely many maximal torsion cosets as such subvarieties are necessarily optimal. We recover the conclusion of the Manin-Mumford Conjecture. Of course in this special case, our argument does not differ significantly from the Pila-Zannier approach [36]. But part (i) of our theorem applied to $s = 0$ yields the following strengthening.

Corollary 9.11. *Let X be an abelian variety defined over a number field K and let $V \subseteq X$ be a subvariety defined over K. Then*

$$\{A \in \mathrm{Opt}(V);\ \langle A\rangle = \langle A\rangle_{\mathrm{geo}}\}$$

is finite.

The next case is $s = 1$; the corresponding case of the Zilber-Pink Conjecture concerns subvarieties of codimension 2. So say $\{P\} \subseteq V$ is an optimal singleton with $\dim\langle P\rangle \leq 1$. If $\dim\langle P\rangle = 0$, then P is of finite order and we are back in the case $s = 0$. So we assume $\dim\langle P\rangle = 1$. We know that $\{P\}$ is geodesic-optimal for V by Proposition 4.5. In other words, P is not contained in a coset of positive dimension contained completely in V. In this setting it would be interesting to know if $\hat{h}(P)$ is bounded from above in terms of V only. The analogous statement in the context of algebraic tori was proved by Bombieri and Zannier, cf. Theorem 1 [51]. Moreover, Checcoli, Veneziano, and Viada [14] showed a related result inside a product of elliptic curves with complex multiplication.

Proof of Theorem 9.8. We can almost prove parts (i) and (ii) simultaneously. However, at times we will branch off the main argument to specialise to the two statements. The proof will be by induction on $\dim V$. Our theorem is trivial if V is a point. Say $\dim V \geq 1$. After replacing K by a finite extension we may suppose that all abelian subvarieties of X and all relevant homomorphisms below are defined over K. We may assume that $\overline{K}$ is a subfield of $\mathbb{C}$. We fix $\mathcal{L}$ an ample, symmetric line bundle on X to make sense of the complexity $\Delta(\cdot)$.

Suppose A is an element of $\mathrm{Opt}(V)$ or $\mathrm{Opt}(V;S)$ depending on whether we are in case (i) or (ii) of the theorem.

In case (i) we suppose $\mathrm{codim}_V A \le r$ and $\dim\langle A\rangle - \dim\langle A\rangle_{\mathrm{geo}} \le s$; in case (ii) we suppose that A contains a point of height at most $(2[K(A):K])^S$.

By Proposition 4.5 the subvariety A is geodesic-optimal for V and thus an irreducible component of $V \cap \langle A\rangle_{\mathrm{geo}}$. By Proposition 6.1 the coset $\langle A\rangle_{\mathrm{geo}}$ is the translate of an abelian subvariety $Y \subseteq X$ that comes from a finite set depending only on V. We observe that this finiteness statement is trivial if $\dim V = 1$; we do not require Proposition 6.1 if V is a curve. We will also fix an ample and symmetric line bundle on the abelian variety X/Y in order to speak of the Néron-Tate height $\hat{h}$.

Let $\varphi : X \to X/Y$ be the canonical morphism. As we are in characteristic 0, there is a Zariski open and dense subset $V' \subseteq V$ such that $\varphi|_{V'} : V' \to \overline{\varphi(V')}$ is a smooth morphism of relative dimension n and $\varphi(V')$ is Zariski open in $\overline{\varphi(V')}$, cf. Corollary III.10.7 [22].

If $A \cap V' = \emptyset$, then A is contained in an irreducible component of $V \smallsetminus V'$ and A is an optimal subvariety for this irreducible component. In both case (i) and (ii) we may apply induction on the dimension as $\dim(V \smallsetminus V') < \dim V$; for case (i) we observe that the codimension in (9.4) drops. So there are only finitely many possibilities for A.

Let us assume $A \cap V' \ne \emptyset$. We note that $A \cap V'$ is an irreducible component of a fibre of $\varphi|_{V'}$. The fibres of $\varphi|_{V'}$ are equidimensional of dimension n. So $\dim A = n$ and φ maps A to some $P \in (X/Y)(\mathbb{C})$.

Since P lies in the torsion coset $\varphi(\langle A\rangle)$ we find $\langle P\rangle \subseteq \varphi(\langle A\rangle)$ and thus $\varphi^{-1}(\langle P\rangle) \subseteq \varphi^{-1}(\varphi(\langle A\rangle)) \subseteq \langle A\rangle + Y$. But Y is contained in a translate of $\langle A\rangle$ and thus $\langle A\rangle + Y = \langle A\rangle$. We conclude

$$\dim Y + \dim\langle P\rangle = \dim \varphi^{-1}(\langle P\rangle) \le \dim\langle A\rangle. \tag{9.5}$$

Next we claim that the singleton $\{P\}$ is an optimal subvariety for $\varphi(V)$. If the contrary holds there is a subvariety B' of $\varphi(V)$ containing P, with positive dimension, and defect at most $\dim\langle P\rangle$. We fix an irreducible component B, that meets V', of the pre-image of B' under $\varphi|_V$ with $A \subseteq B$. As $\varphi|_{V'}$ is smooth of relative dimension $n = \dim A$ we have

$$\dim B = \dim B' + \dim A > \dim A. \tag{9.6}$$

We remark $\langle B\rangle \subseteq \varphi^{-1}(\langle B'\rangle)$, so $\dim\langle B\rangle \le \dim Y + \dim\langle B'\rangle$. Since $\delta(B') \le \dim\langle P\rangle$ we find

$$\dim\langle B\rangle \le \dim Y + \dim B' + \dim\langle P\rangle.$$

Optimality of A and $B \supsetneq A$, a consequence of (9.6), imply

$$\dim\langle A\rangle < \dim A + \dim\langle B\rangle - \dim B \le \dim A + \dim Y + \dim B' + \dim\langle P\rangle - \dim B.$$

We use the equality in (9.6) to find $\dim\langle A\rangle < \dim Y + \dim\langle P\rangle$. This contradicts (9.5) and so $\{P\}$ must be an optimal subvariety for $\varphi(V)$.

Let us suppose $\dim A > 0$ for the moment. Then $\dim \varphi(V) = \dim V - \dim A < \dim V$.

We first branch into case (i). Any singleton in $\varphi(V)$ has codimension $\dim\varphi(V) = \mathrm{codim}_V A \le r$. The bound (9.5) and $\dim Y = \dim\langle A\rangle_{\mathrm{geo}}$ together yield $\dim\langle P\rangle \le \dim\langle A\rangle - \dim\langle A\rangle_{\mathrm{geo}}$. So $\dim\langle P\rangle - \dim\langle P\rangle_{\mathrm{geo}} = \dim\langle P\rangle \le s$ by (9.4). As $\dim\varphi(V) < \dim V$ there are only finitely many possibilities for P by induction. Recall that

$\varphi|_{V'}^{-1}(P)$ contains $A \cap V'$ as an irreducible component. This leaves at most finitely many possibilities for A.

In case (ii) we also want to use induction, but doing so requires a control of the height. By definition there exists $P' \in A(\overline{K})$ with $\hat{h}(P') \leq (2[K(A):K])^S$. Recall that φ comes from a finite set depending only on V. Now $P = \varphi(P')$ and by properties of the Néron-Tate height we have

$$\hat{h}(P) = \hat{h}(\varphi(P')) \leq c_1 \hat{h}(P') \leq c_1(2[K(A):K])^S \tag{9.7}$$

here and below $c_1, c_2, \ldots$ are positive constants that depend only on X, V, and $\mathcal{L}$ but not on A, P, or P'. By Bertrand's Theorem, which we already used in the proof of Lemma 9.3, there exists an abelian subvariety $Z \subseteq X$ such that $\varphi|_Z : Z \to X/Y$ is an isogeny of degree at most c_2. As φ is defined over K we have $[K(P''):K] \leq c_2[K(P):K]$ for any $P'' \in Z$ with $\varphi(P'') = P$. The intersection $V \cap (P' + Y) = V \cap (P'' + Y)$ contains A as an irreducible component. Say $\sigma \in \mathrm{Gal}(\overline{K}/K)$, then $\sigma(A)$ is an irreducible component of $V \cap (\sigma(P'') + Y)$. By Bézout's Theorem the number of irreducible components of this intersection is at most a constant depending only on V and Y. So we have $[K(A):K] \leq c_6[K(P''):K] \leq c_2c_6[K(P):K]$. Inequality (9.7) yields

$$\hat{h}(P) \leq c_1(2c_2c_6[K(P):K])^S.$$

So $\{P\} \in \mathrm{Opt}(\varphi(V); S')$ for an appropriately chosen S'. As in (i) we conclude by induction that A is in a finite set depending only on V.

It remains to treat the case $\dim A = 0$. Then $Y = 0$, φ is the identity, and A consists only of P. Thus $\{P\} \subseteq V$ is an optimal singleton.

In case (i) we first note $\dim V = \mathrm{codim}_V\{P\} \leq r$. So $LGO_s(V)$ holds as X satisfies LGO_s^r. We get

$$\Delta(\langle P \rangle) \leq (2[K(P):K])^\kappa \tag{9.8}$$

where $\kappa > 0$ depends only on X because $\dim\langle P \rangle = \dim\langle P \rangle - \dim\langle P \rangle_{\mathrm{geo}} \leq s$.

In case (ii) we note that $\{P\} \in \mathrm{Opt}(V; S)$ implies the height bound $\hat{h}(P) \leq (2[K(P):K])^S$. We use this bound in connection with Proposition 9.1 and arrive again at an inequality of the form (9.8).

So in any case, we have found a lower bound for the size of the Galois orbit of P. Next we set the stage for the o-minimal machinery. All choices in the following paragraphs are made independently of P unless stated otherwise.

Let us fix a basis $\omega_1, \ldots, \omega_{2g}$ of the period lattice $\Omega_X \subseteq T_0(X)$ as in Section 6.1, here $g = \dim X$. We use it to identify $T_0(X)$ with $\mathbb{R}^{2g}$ as an $\mathbb{R}$-vector space. In these coordinates, $\exp : \mathbb{R}^{2g} \to X(\mathbb{C})$ is a real-analytic group homomorphism with kernel $\mathbb{Z}^{2g}$.

By the discussion above the set

$$F = \{(\psi, w, z) \in \mathrm{Mat}_{2g}(\mathbb{R}) \times \mathbb{R}^{2g} \times \mathbb{R}^{2g};\ z \in \exp|_{[0,1)^{2g}}^{-1}(V(\mathbb{C})) \text{ and } \psi(z - w) = 0\}$$

is definable in $\mathbb{R}_{\mathrm{an}}$ where we identify $\mathrm{Mat}_{2g}(\mathbb{R})$ with $\mathbb{R}^{(2g)^2}$. We will consider F as a definable family parametrised by $\mathrm{Mat}_{2g}(\mathbb{R})$. The kernel of each matrix in $\mathrm{Mat}_{2g}(\mathbb{R})$ is a vector subspace of $\mathbb{R}^{2g}$. In our application, the kernel will be the tangent space of the abelian subvariety determined by $\langle P \rangle$.

Indeed, let us write $\langle P \rangle = Q + Z$ with Z an abelian subvariety of X and where Q has minimal finite order N, i.e. $\Delta(\langle P \rangle) = \max\{N, \deg_{\mathcal{L}} Z\}$. As opposed to Y, we do not yet know that Z comes from a finite set; so we must keep in mind that Q and Z depend on P.

Let $\sigma \in \mathrm{Gal}(\overline{K}/K)$, then $\sigma(P) \in \sigma(Q) + Z$. We write $\sigma(P) = \exp(z_\sigma)$ and $\sigma(Q) = \exp(q_\sigma)$ with $z_\sigma, q_\sigma \in [0,1)^{2g}$. Now $\exp(z_\sigma - q_\sigma) = \sigma(P-Q) \in Z$ implies $z_\sigma - q_\sigma \in \exp^{-1}(Z(\mathbb{C})) = \Omega_X + T_0(Z)$.

Let $\|\cdot\|$ be a norm on $T_0(X)$ as in Section 3.1. According to Lemma 3.2(ii) there exists $\omega_\sigma \in \mathbb{Z}^{2g}$ with $z_\sigma - q_\sigma - \omega_\sigma \in T_0(Z)$ and

$$\|\omega_\sigma\| \le \|z_\sigma - q_\sigma\| + c_4 \deg_{\mathcal{L}} Z.$$

But $\|z_\sigma\| \le c_5$ and $\|q_\sigma\| \le c_5$ as these elements are in the bounded set $[0,1)^{2g}$. Thus $\|\omega_\sigma\| \le c_6 \deg_{\mathcal{L}} Z$. Now ω_σ is integral, hence $H(\omega_\sigma) \le c_7 \deg_{\mathcal{L}} Z$ where the height is as in Section 7 and $c_7 \ge 1$.

As $\sigma(Q)$ has order N we find $q_\sigma \in \frac{1}{N}\mathbb{Z}^{2g}$. The coordinates of q_σ lie in $[0,1)$ and so $H(q_\sigma) \le N$.

Basic height properties yield $H(q_\sigma + \omega_\sigma) \le 2H(q_\sigma)H(\omega_\sigma)$. So

$$H(q_\sigma + \omega_\sigma) \le 2c_7 N \deg_{\mathcal{L}} Z \le 2c_7 \Delta(\langle P\rangle)^2.$$

The tangent space $T_0(Z) \subseteq T_0(X)$ is the kernel of some $\psi \in \mathrm{Mat}_{2g}(\mathbb{R})$. By construction, $(q_\sigma + \omega_\sigma, z_\sigma)$ lies on the fibre F_ψ. The number of distinct z_σ is $[K(P):K]$ as σ ranges over the Galois group. This degree is bounded from below by (9.8). The $q_\sigma + \omega_\sigma$ are rational with height at most $T = 2c_7\Delta(\langle P\rangle)^2 \ge 1$.

There are only finitely many torsion cosets B for with $\Delta(B)$ is bounded by a constant. The singleton $\{P\}$, being optimal for V, is an irreducible component of $V \cap \langle P\rangle$. As our aim is to prove that there are only finitely many P we may assume that $\Delta(\langle P\rangle)$ is sufficiently large with respect to the fixed data. Under this hypothesis and with for example $\epsilon = 1/(4\kappa)$ we can apply Corollary 7.2. We proceed to show that this leads to a contradiction.

There is $\beta : [0,1] \to F_\psi$ as in Corollary 7.2 with Σ the set $\{(q_\sigma + \omega_\sigma, z_\sigma); \sigma \in \mathrm{Gal}(\overline{K}/K)\}$. The o-minimal structure $\mathbb{R}_{\mathrm{an}}$ admits analytic cell decomposition by a result of van den Dries and Miller [17]. So we may assume that β is real analytic on $(0,1)$. The first projection $\beta_1 = \pi_1 \circ \beta : [0,1] \to T_0(X)$ is semi-algebraic and the second one $\beta_2 = \pi_2 \circ \beta : [0,1] \to T_0(X)$ is non-constant. The path β_2 begins at $\beta_2(0) = z_\sigma$ for some $\sigma \in \mathrm{Gal}(\overline{K}/K)$. The image of $\beta_2 - \beta_1$ lies in $\ker\psi = T_0(Z)$ by our choice of ψ. So $\phi \circ \exp \circ \beta_1 = \phi \circ \exp \circ \beta_2$ where $\phi : X \to X/Z$ is the quotient morphism.

We claim that $\phi \circ \exp \circ \beta_1$ is non-constant. Let us assume the contrary, then $\phi \circ \exp \circ \beta_2$ is constant too. As $\exp \circ \beta_2(0) = \sigma(P)$ we have $(\exp \circ \beta_2)([0,1]) \subseteq \sigma(P) + Z = \sigma(\langle P\rangle)$. But this image of $[0,1]$ lies in $V(\mathbb{C})$ by the definition of F. So $(\exp \circ \beta_2)([0,1]) \subseteq V \cap \sigma(\langle P\rangle) = \sigma(V \cap \langle P\rangle)$. Recall that $\{P\}$ is an optimal singleton for V. So P is isolated in $V \cap \langle P\rangle$ and thus $\sigma(P)$ is isolated in $\sigma(V \cap \langle P\rangle)$. This contradicts the fact that β_2 is continuous and non-constant.

Let $R \subseteq X(\mathbb{C})$ denote the image of $\exp \circ \beta_1$, it is an uncountable set by the previous paragraph. The differential $d\phi : T_0(X) \to T_0(X/Z)$ of ϕ is a linear map. So $\phi(R)$ is the image of the semi-algebraic map $d\phi \circ \beta_1$ composed with $T_0(X/Z) \to (X/Z)(\mathbb{C})$. The Ax-Lindemann-Weierstrass Theorem 5.4 implies that $\phi(\mathrm{Zcl}(R)) = \mathrm{Zcl}(\phi(R)) \subseteq X/Z$ is a positive dimensional coset. We abbreviate $C = \sigma^{-1}(\mathrm{Zcl}(R))$ which contains P as a point. Then C must be irreducible, as $\beta_1|_{(0,1)}$ is real analytic, and of positive dimension. The image $\phi(C)$ is a translate of Z'/Z where $Z' \subseteq X$ is an abelian subvariety that contains Z. Now C is contained in the coset $\phi^{-1}(\phi(C)) =$

$P + Z'$. This is even a torsion coset because $P + Z' \supseteq P + Z = \langle P\rangle$ contains a point of finite order. Therefore,

$$\langle C\rangle \subseteq P + Z'.$$

Basic dimension theory yields the inequality in $\dim Z'/Z = \phi(C) \le \dim C$. Thus

$$\delta(C) = \dim\langle C\rangle - \dim C \le \dim(P+Z') - \dim C \le \dim Z' - \dim Z'/Z = \dim Z = \delta(P).$$

But $\dim C \ge 1$ and this contradicts the optimality of $\{P\}$. □

9.3. **Intersecting with algebraic subgroups.** In this section we prove that a height upper bound for curves due to Rémond in combination with the o-minimal machinary is strong enough to establish LGO_s^1 for all abelian varieties and all $s \ge 0$. In turn this will yield Theorem 1.1, the Zilber-Pink Conjecture for curves in abelian varieties when all geometric objects are defined over an algebraic closure of the rationals. We also prove some partial results in the direction of this conjecture for higher dimensional subvarieties.

Theorem 9.12 (Rémond)**.** *Let X be an abelian variety defined over a number field K, equipped with an ample, symmetric line bundle and its associated Néron-Tate height. Suppose that V is a curve in X that is not contained in any proper abelian subvariety of X. Then the Néron-Tate height is bounded from above on $V(\overline{K}) \cap X^{[2]}$.*

Proof. This is Rémond's corollaire 1.6 [43]. □

Corollary 9.13. *Any abelian variety defined over a number field satisfies LGO_s^1 for all $s \ge 0$.*

Proof. Let $V \subseteq X$ be a subvariety with $\dim V \le 1$. We may clearly assume that V is a curve. If $\{P\} \subseteq V(\overline{K})$ is an optimal singleton, then $\dim\langle P\rangle = \delta(\{P\}) < \delta(V) = \dim\langle V\rangle - 1$. After translating V by a torsion point it is contained in the abelian subvariety of X determined by $\langle V\rangle$. Without loss of generality it suffices to verify LGO_s^1 with X replaced by this abelian subvariety. Now $\langle V\rangle = X$ and P is contained in an abelian subvariety of X of codimension at least 2. So the height of P is bounded from above by Rémond's Theorem. On inserting this height bound in Proposition 9.1 we find that $LGO_s(V)$ is satisfied. □

We combine the corollary with results from the previous section to obtain the following strengthening of Theorem 1.1.

Theorem 9.14. *Let X be an abelian variety defined over a number field K. Suppose $V \subseteq X$ is a subvariety defined over K.*

(i) *The set*
$$\{A \in \mathrm{Opt}(V);\ \mathrm{codim}_V A \le 1\}$$
is finite.

(ii) *If V is a curve then* $\mathrm{Opt}(V)$ *is finite.*

(iii) *If V is a curve that is not contained in a proper abelian subvariety of X, then $V(\overline{K}) \cap X^{[2]}$ is finite.*

Proof. Part (i) follows from the corollary above and from Theorem 9.8(i). Part (ii) is a special case of (i) and to see (iii) we note that any point in the intersection defines an optimal singleton for V. □

The next theorem improves on Theorem 1.3 stated in the introduction. The open anomalous set V^{oa} was introduced in Definition 6.4, we use it in part (iii) below. Parts (iii) and (iv) rely on a height bound of the first-named author whereas parts (v) and (vi) use a result of Rémond.

Theorem 9.15. *Let X be an abelian variety defined over a number field K, equipped with an ample, symmetric line bundle and its associated Néron-Tate height $\hat{h}$. Suppose $V\subseteq X$ is a subvariety defined over K.*

(i) *If $S\geq 0$ then*
$$\left\{P\in V(\overline{K})\cap X^{[1+\mathrm{codim}_X\langle V\rangle+\dim V]};\ \hat{h}(P)\leq S\right\}$$
is not Zariski dense in V.

(ii) *If $S\geq 0$ then*
$$\left\{P\in V(\overline{K})\cap X^{[1+\dim V]};\ \hat{h}(P)\leq S\right\}$$
is contained in a finite union of proper algebraic subgroups of X.

(iii) *The set $V^{\mathrm{oa}}(\overline{K})\cap X^{[1+\dim V]}$ is finite.*

(iv) *Suppose $\dim V\geq 1$ and $\dim\varphi(V)=\min\{\dim X/Y,\dim V\}$ for all abelian subvarieties $Y\subseteq X$ where $\varphi: X\to X/Y$ is the canonical morphism. Then $V(\overline{K})\cap X^{[1+\dim V]}$ is not Zariski dense in V.*

(v) *Let $\Gamma\subseteq X(\overline{K})$ be a finitely generated subgroup and*
$$\overline{\Gamma}=\{P\in X(\overline{K});\ \text{there is an integer } n\geq 1 \text{ with } nP\in\Gamma\}.$$
Then
$$\bigcup_{\substack{H\subseteq X\\ H\ \text{algebraic subgroup}\\ \mathrm{codim}_X H\geq 1+\dim V}} V^{\mathrm{oa}}(\overline{K})\cap(H+\overline{\Gamma}). \tag{9.9}$$
is finite.

(vi) *Let V be as in (iv) and $\overline{\Gamma}$ be as in (v). Then*
$$\bigcup_{\substack{H\subseteq X\\ H\ \text{algebraic subgroup}\\ \mathrm{codim}_X H\geq 1+\dim V}} V(\overline{K})\cap(H+\overline{\Gamma})$$
is not Zariski dense in V.

Proof. For part (i) let $P\in V(\overline{K})\cap X^{[1+\mathrm{codim}_X\langle V\rangle+\dim V]}$ with $\hat{h}(P)\leq S$. We can enlargen $\{P\}$ to an optimal subvariety A for V with $\delta(A)\leq\dim\langle P\rangle$. As A contains a point of height at most $S\leq(2[K(A):K])^S$ we find $A\in\mathrm{Opt}(V;S)$. But the latter set is finite according to Theorem 9.8(ii). To prove (i) it suffices to establish $A\neq V$. This follows from $\delta(A)\leq\dim\langle P\rangle<\dim\langle V\rangle-\dim V$.

Part (ii) is a consequence of (i). Indeed, there is nothing to show if V is contained in a proper algebraic subgroup of X or if V is a point. Otherwise we have $\langle V\rangle=X$ and by (i) the set in the assertion is contained in $V_1\cup\cdots\cup V_n$ where $V_i\subsetneq V$ are subvarieties of V. We conclude (ii) by induction on $\dim V>\dim V_i$.

To show (iii) we may assume $V\neq X$ and that V is not contained in a proper abelian subvariety of X. Indeed, if the second condition fails, then $V^{\mathrm{oa}}=\emptyset$. We know from the main result of [20] that the Néron-Tate height is bounded from above by S, say, on $V^{\mathrm{oa}}(\overline{K})\cap X^{[\dim V]}$ and thus in particular on $V^{\mathrm{oa}}(\overline{K})\cap X^{[1+\dim V]}$.

Rémond [43, 44] independently obtained a height bound for the latter set, in his notation we have $V \setminus V^{\mathrm{oa}} = Z^{(1+\dim V)}_{V,\mathrm{an}}$. Say P is in the intersection in part (iii). If A is an optimal subvariety for V containing P with $\delta(A) \leq \dim\langle P\rangle$ then $\dim X - \dim V - 1 \geq \dim\langle P\rangle \geq \dim\langle A\rangle - \dim A$. But $P \in V^{\mathrm{oa}}(\overline{K})$ entails $\dim A = 0$. So $A = \{P\} \in \mathrm{Opt}(V;S)$ and the (iii) follows as (i) because $\mathrm{Opt}(V;S)$ is finite.

We recall that V^{oa} is Zariski open in V. Part (iv) follows from (iii) since the condition on V entails $V^{\mathrm{oa}} \neq \emptyset$ by the final comment in Section 6.1.

For part (v) we will require Rémond's height bound in his Théorème 1.2 [43] combined with his Théorème 1.4 [44]. Together they imply that there is an upper bound for the Néron-Tate height of the points in $V^{\mathrm{oa}}(\overline{K}) \cap (\overline{\Gamma} + X^{[1+\dim V]})$.

We proceed by following the argumentation in the proof of Pink's Theorem 5.3 [38] which we briefly sketch. Suppose $P_1, \ldots, P_t$ are $\mathbb{Z}$-independent elements that generate a subgroup of finite index in Γ. After replacing $(P_1, \ldots, P_t) \in X^t(\overline{K})$ by a positive multiple, we may assume that this point generates a subgroup of $X^t(\overline{K})$ whose Zariski closure in X^t is an abelian subvariety Y. Let $V' = V \times \{(P_1, \ldots, P_t)\}$, this is a subvariety of X^{t+1}. Now any point P in (9.9) is a $\mathbb{Z}$-linear combination of the P_i modulo an algebraic subgroup of codimension at least $\dim V' + 1$. So the augmented point $P' = (P, P_1, \ldots, P_t)$ lies in $V'(\overline{K}) \cap (X \times Y)^{[\dim V'+1]}$ because $\dim V' = \dim V$. The Néron-Tate height of P' is bounded by a constant S that only depends on V and the P_i.

We proceed as in the end of (iii). Let A' be an optimal subvariety for V' that contains the point P' and with $\delta(A') \leq \delta(\{P'\})$. So $\dim\langle A'\rangle - \dim A' \leq \dim\langle P'\rangle \leq \dim X \times Y - \dim V' - 1$. The projection of $\langle A'\rangle \subseteq X \times Y$ to Y is an irreducible component of an algebraic group which contains $(P_1, \ldots, P_t)$; so it must equal Y. Therefore, each fibre of this projection is a coset of dimension $\dim\langle A'\rangle - \dim Y$. We observe that $A' = A \times \{(P_1, \ldots, P_t)\}$ is contained in such a fibre, thus

$$\dim\langle A\rangle_{\mathrm{geo}} \leq \dim\langle A'\rangle - \dim Y \leq \dim A' + \dim X - \dim V' - 1 = \dim A + \dim X - \dim V - 1$$

and hence

$$\dim A \geq \dim V + \dim\langle A\rangle_{\mathrm{geo}} - \dim X + 1.$$

Recall that $P \in A(\overline{K})$; we must have $\dim A = 0$ because $P \in V^{\mathrm{oa}}(\overline{K})$. Thus $A' = \{P'\}$. Part (v) follows as A' lies in the set $\mathrm{Opt}(V';S)$ which is finite by Theorem 9.8(ii).

The claim in (vi) follows from (v) as V^{oa} is Zariski open in V and non-empty. □

The case of curves. We give a brief sketch of how the argument plays out for curves, as several aspects become simpler.

Sketch proof of Theorem 1.1. We consider a curve V contained in an abelian variety X of dimension g, both defined over a number field K. We suppose that V is not contained in any proper special subvariety of X. Then an optimal proper subvariety of V is a point which lies in a special subvariety of X of codimension ≥ 2.

By Corollary 9.13, using the result 9.12 of Rémond, X satisfies LGO^1_s, so if $P \in V(\overline{K})$ with $\{P\}$ optimal then $\Delta(\langle P\rangle) \leq (2[K(P):K])^{\kappa}$ for a suitable positive κ.

Now the intersection of V with a codimension 2 special subvariety gives rise to a rational point on the projection to the middle factor $\mathbb{R}^{2g}$ of the definable set

$$\{(\psi, w, z) \in M \times \mathbb{R}^{2g} \times \mathbb{R}^{2g} : z \in \exp|^{-1}_{[0,1)^{2g}}(V(\mathbb{C})) \text{ and } \psi(z-w) = 0\}$$

where ψ is in a real semi-algebraic set M parameterising the subspaces of $\mathbb{C}^g$ of complex codimension 2, and $w \in \mathbb{R}^{2g}$ parameterises translations of these subspaces.

The height in the space of translations $\mathbb{R}^{2g}$ of the point corresponding to $P \in \langle P \rangle$ is $\ll \Delta(\langle P \rangle)^2$, and so P and its conjugates give rise to "many" rational points on $\mathbb{R}^{2g}$.

If $\Delta(\langle P \rangle)$ is sufficiently large, then Corollary 7.2 implies that there is a semi-algebraic curve in $\mathbb{R}^{2g}$, parameterised by some real parameter t say, for which the corresponding linear varieties have a non-constant intersection with $\exp|_{[0,1)^{2g}}^{-1}(V(\mathbb{C}))$. Here we use the fact that all estimates are uniform over M. The linear varieties are all translates of the tangent space of a single abelian subvariety Z of X with $\mathrm{codim}_X Z \geq 2$. The point P lies in the translate of Z by a point of finite order. Composing the said semi-algebraic curve with $\mathbb{R}^{2g} \to X(\mathbb{C}) \to (X/Z)(\mathbb{C})$ gives us a non-constant mapping that can be analysed using Ax-Schnauel in the guise of Theorem 5.3. The image of this semi-algebraic curve lies in $V + Z$, the image of V under the quotient map $X \to X/Z$. So $V + Z$ is a coset. But it contains $P + Z$, a point of finite order, and so it is even a torsion coset.

The torsion coset has dimension 1 and this contradicts the assumption on V as $\dim V = 1 < \dim X/Z$. Hence $\Delta(\langle P \rangle)$ is bounded for optimal P, and hence only finitely many special subvarieties (and hence optimal points) may arise. □

10. Unlikely intersections in $Y(1)^n$

Theorem 10.1. *Assume LGO and WCA for $Y(1)^n$. Let $X \subseteq Y(1)^n$ be a special subvariety and $V \subseteq X$. Then $\mathrm{Opt}(V)$ is a finite set.*

Proof. We prove the theorem by induction on V, the case $\dim V = 0$ being trivial. So we assume that $\dim V \geq 1$ and the theorem holds for all V of smaller dimension. Let K be a field of definition for V which is finitely generated over $\mathbb{Q}$.

By Proposition 4.5, an optimal component is geodesic-optimal. By Proposition 6.6 the set of "basic special subvarieties" that have a translate which is geodesic-optimal is finite. So the subvarieties comprising $\mathrm{Opt}(V)$ are components of the intersection of V with the translates of finitely many basic special subvarieties $T \subseteq X$. One such T may of course be the whole of $Y(1)^n$, with X_T being also $Y(1)^n$ parameterising individual points of $Y(1)^n$.

Fix such T. It evidently suffices now to show that only finitely many translates of T are such that $V \cap T$ has components which are optimal. Let X_T denote the translate space of T, which is a suitable $Y(1)^m$.

We have a "quotient map" $\phi : T \to Y(1)^m$, the fibre over $t \in Y(1)^m$ being T_t. Write τ for the dimension of these fibres. By [22, Corollary III.10.7], as we are in characteristic 0, there is a Zariski-open (in V) and dense $V' \subseteq V$ on which the restriction $\phi|_{V'} : V' \to \phi(V')$ is a smooth morphism of relative dimension ν. We may further assume that $\phi(V')$ is Zariski open in its Zariski closure, which we denote $V_T \subseteq X_T$.

Now suppose C is an optimal component of dimension d and geodesic defect δ, a component of $V \cap T_y$ for some y. If $C \cap V' = \emptyset$ then C is contained in an irreducible component of $V - V'$, and is an optimal component for this irreducible component. As this irreducible component has lower dimension than V we conclude by the induction on $\dim V$ that there are only finitely many such C.

So we may suppose $C \cap V' \neq \emptyset$. Since $C \cap V'$ is an irreducible component of a fibre of $\phi|_{V'}$ we have $n = \dim C$; also the image of C in V_T under ϕ is the point y.

We further observe that if $y \in A \subseteq V_T$ is contained in a special subvariety $S \subseteq X_T$ then $\phi^{-1}(A)$ is contained in a special subvariety $\phi^{-1}(S)$ of dimension $\dim S + \tau$. If we take A to be a component of $\phi|_{V'}^{-1}$ containing C we see that

$$\delta(A) \leq \dim S + \tau - \dim A = \dim S + \tau - (\dim A' + \dim C).$$

Next we claim that $\{y\}$ is an optimal subvariety for V_T. Note that $\phi(\langle C\rangle)$ is special of dimension $\dim C + \delta(C) - \tau$ and contains y. Now suppose that A is a component of V_T with $\{y\} \subseteq A$ and

$$\dim\langle A\rangle - \dim A \leq \dim C + \delta(C) - \tau.$$

Let B be the component of $\phi^{-1}(A)$ containing C. Then

$$\dim B = \dim\left(\phi^{-1}(A) \cap V'\right) = \dim A + \nu$$

and

$$\delta(B) \leq \dim \phi^{-1}\langle A\rangle - \dim B \leq \dim A + \dim C + \delta(C) - (\dim A + \nu) = \delta(C)$$

and so by optimality of C we must have $B = C$ and $A = \{y\}$.

By induction, if $\dim V_T < \dim V$ then V_T has only finitely many optimal subvarieties. We are reduced to the case $\dim V_T = \dim V$, which is the case that T is the family of points. We take a finite extension field L of K over which all optimal subvarieties of positive dimension are defined.

Now suppose that there is a special subvariety S intersecting V optimally in a point $\{y\}$. Let M^R be the family of Mobius subvarieties containing the components of $\pi^{-1}(S)$ (note that the family of all Mobius subvarieties is definable). Let $Z \subseteq M_R \times Y(1)^n$ be the (definable) set of pairs (t, u) such that $\{u\}$ is an optimal component of $V \cap \pi(M_t^R \cap \mathbb{F}_0^n)$. Let κ be the constant afforded by LGO for V. We apply Corollary 7.2 (with $\ell = 0$) and $\epsilon = (20\kappa)^{-1}$ to get $c = c(Z, 2, \epsilon)$. Let $\Sigma = Z^{\sim}(2, T)$.

Let $T \geq 1$ and suppose $\#\pi_2(\Sigma) > cT^\epsilon$. Then we have a curve β in one of the constituent definable blocks such that $\pi_1 \circ \beta$ is semialgebraic and $\pi_2 \circ \beta$ is non-constant. The union of the Mobius subvarieties over the complexification of $\pi_1 \circ \beta$ meets $\pi^{-1}(V)$ in an uncountable set, hence in a set of complex dimension at least one. This union of Mobius varieties is thus a larger algebraic subvariety of $\mathbb{H}^n$ with the same defect (in the sense of 5.11) as each fibre. By WCA there is a geodesic component containing this subvariety of the same defect. The weakly special subvariety is however special, as it contains some of the special subvarieties (conjugates of S) corresponding to $\{y\}$ and its conjugates. This contradicts the assumption that $\{y\}$ and its conjugates are optimal.

Therefore $\#\pi_2(\Sigma) \leq cT^\epsilon$. But if $T = \Delta(\langle y\rangle)^{10}$ we have $\#\pi_2(\Sigma) \geq T^{2\epsilon}/(2[L : K])$ by LGO and Proposition 6.8. Therefore $\Delta(\langle y\rangle)$ is bounded. This completes the proof of Theorem 10.1. □

Acknowledgements

The authors thank Martin Bays and Martin Hils for informing them about the work of Poizat and Fabien Pazuki for comments on the text. PH was partially supported by the National Science Foundation under agreement No. DMS-1128155. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. JP acknowledges with thanks that his research was supported in part by a grant from the EPSRC entitled "O-minimality and diophantine geometry", reference EP/J019232/1.

References

1. F. Amoroso and S. David, *Le problème de Lehmer en dimension supérieure*, J. Reine Angew. Math. **513** (1999), 145–179.
2. J. Ax, *On Schanuel's conjectures*, Ann. of Math. (2) **93** (1971), 252–268.
3. ______, *Some topics in differential algebraic geometry I: Analytic subgroups of algebraic groups*, Amer. J. Math. **94** (1972), 1195–1204.
4. A. Baudisch, M. Hils, A. Martin-Pizarro, and F. Wagner, *Die böse Farbe,* J. I. M. Jussieu **8** (2007), 415–443.
5. D. Betrand and V. Zudilin, *Derivatives of Siegel modular forms, and exponential functions*, Izv. Math. **65** (2001), no. 4, 659 – 671.
6. C. Birkenhake and H. Lange, *Complex Abelian Varieties*, Springer, 2004.
7. E. Bombieri and W. Gubler, *Heights in Diophantine Geometry*, Cambridge University Press, 2006.
8. E. Bombieri, D. W. Masser, and U. Zannier, *On unlikely intersections of complex varieties with tori*, Acta Arith. **133** (2008), no. 4, 309–323.
9. ______, *Intersecting a curve with algebraic subgroups of multiplicative groups*, Internat. Math. Res. Notices (1999), no. 20, 1119–1140.
10. ______, *Anomalous Subvarieties - Structure Theorems and Applications*, Int. Math. Res. Not. IMRN (2007), no. 19, 1–33.
11. L. Capuano, *Unlikely Intersections and Applications to Diophantine Geometry*, Ph.D. thesis, 2013.
12. M. Carrizosa, *Problème de Lehmer et variétés abéliennes CM*, C. R. Math. Acad. Sci. Paris **346** (2008), no. 23-24, 1219–1224.
13. ______, *Petits points et multiplication complexe*, Int. Math. Res. Not. (2009), no. 16, 3016–3097.
14. S. Checcoli, F. Veneziano, and E. Viada, *On Torsion Anomalous Intersections (with an Appendix by P. Philippon)*, (preprint arxiv 1204.1435).
15. L. van den Dries, *Tame topology and o-minimal structures*, **248** (1998), x+180.
16. L. van den Dries, A. Macintyre, and D. Marker, *The elementary theory of restricted analytic fields with exponentiation*, Ann. of Math. (2) **140** (1994), no. 1, 183–205.
17. L. van den Dries and C. Miller, *On the real exponential field with restricted analytic functions*, Israel J. Math. **85** (1994), no. 1-3, 19–56.
18. A. Galateau, *Une minoration du minimum essentiel sur les variétés abéliennes*, Comment. Math. Helv. **85** (2010), no. 4, 775–812.
19. H. Grauert and R. Remmert, *Coherent Analytic Sheaves*, Springer, 1984.
20. P. Habegger, *Intersecting subvarieties of abelian varieties with algebraic subgroups of complementary dimension*, Invent. Math. **176** (2009), 405–447.
21. P. Habegger and J. Pila, *Some unlikely intersections beyond André-Oort*, Compos. Math. **148** (2012), no. 1, 1–27.
22. R. Hartshorne, *Algebraic Geometry*, Springer, 1997.
23. J. Kirby, *The theory of the exponential differential equations of semiabelian varieties*, Selecta Math. (N.S.) **15** (2009), no. 3, 445–486.
24. S. Lang, *Introduction to transcendental numbers*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1966.
25. ______, *Elliptic functions*, second ed., Graduate Texts in Mathematics, vol. 112, Springer-Verlag, New York, 1987, With an appendix by J. Tate.
26. K. Mahler, *On algebraic differential equations satisfied by automorphic functions.*, J. Austral. Math. Soc. **10** (1969), 445–450.
27. D. W. Masser, *Small values of the quadratic part of the Néron-Tate height on an abelian variety*, Compositio Math. **53** (1984), no. 2, 153–170.
28. D.W. Masser and U. Zannier, *Torsion anomalous points and families of elliptic curves*, C. R. Acad. Sci. Paris, Sér. I **346** (2008), 491–494.
29. G. Maurin, *Courbes algébriques et équations multiplicatives*, Math. Ann. **341** (2008), 789–824.

30. Y. Peterzil and S. Starchenko, *Tame complex analysis and o-minimality*, Proceedings of the International Congress of Mathematicians. Volume II, pp. 58–81.
31. ______, *Uniform definability of the Weierstrass ℘ functions and generalized tori of dimension one*, Selecta Math. (N.S.) **10** (2004), no. 4, 525–550.
32. J. Pila, *O-minimality and the André-Oort conjecture for* $\mathbb{C}^n$, Ann. of Math. (2011), no. 173, 1779–1840.
33. ______, *Special point problems with elliptic modular surfaces*, Mathematika **60** (2014), 1–31.
34. J. Pila, J. Tsimerman, *Ax-Schanuel for the j-function*, (preprint arxiv:1412.8255).
35. J. Pila and A. J. Wilkie, *The rational points of a definable set*, Duke Math. J. **133** (2006), no. 3, 591–616.
36. 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), no. 2, 149–162.
37. R. Pink, *A combination of the conjectures of Mordell-Lang and André-Oort*, Geometric methods in algebra and number theory, Progr. Math., vol. 235, Birkäuser, 2005, pp. 251–282.
38. ______, *A Common Generalization of the Conjectures of André-Oort, Manin-Mumford, and Mordell-Lang*, Preprint (2005), 13pp.
39. B. Poizat, *L'egalité au cube,* J. Symbolic Logic **66** (2001), 1647–1676.
40. N. Ratazzi, *Intersection de courbes et de sous-groupes et problèmes de minoration de dernière hauteur dans les variétés abéliennes C.M*, Ann. Inst. Fourier (Grenoble) **58** (2008), no. 5, 1575–1633.
41. N. Ratazzi and E. Ullmo, *Galois + équidistribution = Manin-Mumford*, Arithmetic geometry, Clay Math. Proc., vol. 8, Amer. Math. Soc., Providence, RI, 2009, pp. 419–430.
42. G. Rémond, *Intersection de sous-groupes et de sous-variétés I*, Math. Ann. **333** (2005), no. 3, 525–548.
43. ______, *Intersection de sous-groupes et de sous-variétés II*, J. Inst. Math. Jussieu **6** (2007), no. 2, 317–348.
44. ______, *Intersection de sous-groupes et de sous-variétés. III*, Comment. Math. Helv. **84** (2009), no. 4, 835–863.
45. G. Rémond and E. Viada, *Problème de Mordell-Lang modulo certaines sous-variétés abéliennes*, Internat. Math. Res. Notices **35** (2003), 1915–1931.
46. J. Tsimerman, *Brauer-Siegel for arithmetic tori and lower bounds for Galois orbits of special points*, J. Amer. Math. Soc. **25** (2012), no. 4, 1091–1117.
47. E. Ullmo, *Quelques applications du théorème de Ax-Lindemann hyperbolique*, (2012 preprint, available from the author's webpage, and *Compositio*, to appear.).
48. E. Ullmo and A. Yafaev, *Nombres de classes des tores de multiplication complexe et bornes inférieures pour orbites Galoisiennes de points spéciaux*, Bull. Soc. Math. France **143** (2015), no. 1, 197–228.
49. E. Viada, *The intersection of a curve with algebraic subgroups in a product of elliptic curves*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **2** (2003), no. 1, 47–75.
50. ______, *The intersection of a curve with a union of translated codimension-two subgroups in a power of an elliptic curve*, Algebra Number Theory **2** (2008), no. 3, 249–298.
51. U. Zannier, *Appendix* by Umberto Zannier, pp. 517-539 in Schinzel, *Polynomials with special regard to reducibility Encylopedia of Mathematics and its Applications* **77***, CUP, 2000.*, 2000.
52. ______, *Some problems of unlikely intersections in arithmetic and geometry*, Annals of Mathematics Studies, vol. 181, Princeton University Press, Princeton, NJ, 2012, With appendixes by David Masser.
53. B. Zilber, *Exponential sums equations and the Schanuel conjecture*, J. London Math. Soc. (2) **65** (2002), no. 1, 27–44.

Departement Mathematik und Informatik
Universität Basel
Spiegelgasse 1
4059 Basel
Switzerland
philipp.habegger@unibas.ch

Mathematical Institute
University of Oxford
Woodstock Road
Oxford OX2 6GG
UK
pila@maths.ox.ac.uk

June 5, 2015