Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XVII (2017), 1357-1382

Multiplicative relations among singular moduli

JONATHAN PILA AND JACOB TSIMERMAN

Abstract. We consider some Diophantine problems of mixed modular-multiplicative type. In particular, we prove, for each $n \geq 1$, a finiteness result for n-tuples of singular moduli minimally satisfying a non-trivial multiplicative relation.

Mathematics Subject Classification (2010): 14G18 (primary); 03C64 (secondary).

1. Introduction

We consider some Diophantine problems of mixed modular-multiplicative type associated with the Zilber-Pink conjecture (ZP; see [4, 26, 35] and Section 2). Our results rely on the "modular Ax-Schanuel" theorem recently established by us [24].

Recall that a *singular modulus* is a complex number which is the j-invariant of an elliptic curve with complex multiplication; equivalently it is a number of the form $\sigma = j(\tau)$ where $j : \mathbb{H} \to \mathbb{C}$ is the elliptic modular function, $\mathbb{H} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ is the complex upper-half plane, and $\tau \in \mathbb{H}$ is a quadratic point (*i.e.* $[\mathbb{Q}(\tau) : \mathbb{Q}] = 2$).

Definition 1.1. An n-tuple $(\sigma_1, \ldots, \sigma_n)$ of distinct singular moduli will be called a *singular-dependent n-tuple* if the set $\{\sigma_1, \ldots, \sigma_n\}$ is multiplicatively dependent (*i.e.* $\prod \sigma_i^{a_i} = 1$ for some integers a_i not all zero), but no proper subset is multiplicatively dependent.

Theorem 1.2. *Let $n \geq 1$. There exist only finitely many singular-dependent n-tuples.*

The independence of proper subsets is clearly needed to avoid trivialities. The result is ineffective. Some examples (including a singular-dependent 5-tuple) can be found among the rational singular moduli (listed in [29, A.4]; see 6.3). Bilu-Masser-Zannier [3] show that there are no singular moduli with $\sigma_1\sigma_2 = 1$. This

JP was supported by an EPSRC grant entitled "O-minimality and Diophantine geometry", reference EP/J019232/1. JT was supported by an NSERC discovery grant.

Received February 3, 2015; accepted in revised form June 14, 2016.
Published online December 2017.

result is generalised by Bilu-Luca-Pizarro-Madariaga [2] to classify all solutions of $\sigma_1\sigma_2 \in \mathbb{Q}^\times$. Habegger [12] shows that only finitely many singular moduli are algebraic units.

In addition to the "modular Ax-Schanuel", we make use of isogeny estimates and other arithmetic ingredients, gathered in Section 6, and we require the following result showing that distinct rational "translates" of the j-function are multiplicatively independent modulo constants. To formulate it, recall that, for $g_1, g_2 \in \mathrm{GL}_2^+(\mathbb{Q})$, the functions $j(g_1 z)$, $j(g_2 z)$ are identically equal if and only if $[g_1] = [g_2]$ in $\mathrm{PSL}_2(\mathbb{Z})\backslash\mathrm{PGL}_2^+(\mathbb{Q})$; functions $f_1, \ldots, f_k : \mathbb{H} \to \mathbb{C}$ will be called *multiplicatively independent modulo constants* if there is no relation $\prod_{i=1}^k f_i^{n_i} = c$ where n_i are integers, not all zero, and $c \in \mathbb{C}$.

Theorem 1.3. *Let $g_1, \ldots, g_k \in \mathrm{GL}_2^+(\mathbb{Q})$. If the functions $j(g_1 z), \ldots, j(g_k z)$ are pairwise distinct then they are multiplicatively independent modulo constants.*

Theorem 1.3 is not predicted by ZP, nor would it follow from "Ax-Schanuel" for exp and j (see Section 3). But in view of Theorem 1.3, Theorem 1.2 is implied by ZP.

The Zilber-Pink setting is introduced in Section 2. After the proofs of 1.3 and 1.2 in Section 4 and Section 6, we discuss further ZP problems in the same setting in Section 7, Section 8, and Section 9, obtaining some partial results and some results conditional on certain "weakly bounded height conjectures" which we formulate in this setting. These are along the lines of a conjecture of Habegger [10] (see also [11, Appendix B]) in the modular setting, itself an analogue of the "Bounded Height Conjecture" for $(\mathbb{C}^\times)^n$ formulated by Bombieri-Masser-Zannier [4] and proved by Habegger [9].

ACKNOWLEDGEMENTS. We thank Gareth Jones for comments. We are grateful to the referee for corrections and careful comments as well as for raising the question addressed in Section 9.

2. The Zilber-Pink setting

We identify varieties and subvarieties with their sets of complex points (thus $Y(1)(\mathbb{C}) = \mathbb{C}$ and $\mathbb{G}_\mathrm{m}(\mathbb{C}) = \mathbb{C}^\times$). Varieties and subvarieties are assumed irreducible over $\mathbb{C}$.

For $m, n \in \mathbb{N} = \{0, 1, 2, \ldots\}$ set

$$X = X_{m,n} = Y(1)^m \times \mathbb{G}_\mathrm{m}^n.$$

Definition 2.1.

1. A *weakly special subvariety* of $Y(1)^m = X_{m,0} = \mathbb{C}^m$ is a subvariety of the following form. There is a "partition" $m_0, \ldots, m_k$ of $\{1, \ldots, m\}$, in which only

m_0 is permitted to be 0, but $k = 0$ is permitted such that $M = M_0 \times M_1 \times \ldots \times M_k$ where M_0 is a point in $\mathbb{C}^{m_0}$ (here $\mathbb{C}^{m_i}$ refers to the cartesian product of the coordinates contained in m_i, which is a subset of $\{1, \ldots, m\}$) and, for $i = 1, \ldots, k$, $M_i \subset \mathbb{C}^{m_i}$ is a modular curve;

2. A *special point* of $\mathbb{C}^m$ is a weakly special subvariety M of dimension zero (so $n_0 = \{1, \ldots, n\}$ and $M = M_0$) such that each coordinate of M is a singular modulus;
3. A *special subvariety* of $\mathbb{C}^m$ is a weakly special subvariety such that $m_0 = \emptyset$ or $M_0 \in \mathbb{C}^{m_0}$ is a special point. It is *strongly* special if $m_0 = \emptyset$;
4. A *weakly special subvariety* of $\mathbb{G}_{\mathrm{m}}^n = X_{0,n} = (\mathbb{C}^\times)^n$ is a coset of a subtorus, *i.e.* a subvariety defined by a finite system of equations $\prod x_i^{a_{ij}} = \xi_j$, $j = 1, \ldots, k$ where, for each j, $a_{ij} \in \mathbb{Z}$ are not all zero, $\xi_j \in \mathbb{C}^\times$ and the lattice generated by the exponent vectors $(a_{1j}, \ldots, a_{nj})$, $j = 1, \ldots, k$ is primitive;
5. A *special point* of $\mathbb{G}_{\mathrm{m}}^n$ is a torsion point;
6. A *special subvariety* of $\mathbb{G}_{\mathrm{m}}^n$ is a weakly special subvariety such that each ξ_j is a root of unity; *i.e.* it is a coset of a subtorus by a torsion point;
7. A *weakly special subvariety* of X is a product $M \times T$ where M, T are weakly special subvarieties of $Y(1)^m$, $\mathbb{G}_{\mathrm{m}}^n$, respectively, and likewise for a *special point* of X and *special subvariety* of X.

Definition 2.2. Let $W \subset X$ be a subvariety. A subvariety $A \subset W$ is called an *atypical component* (of W in X) if there is a special subvariety $T \subset X$ such that $A \subset W \cap T$ and

$$\dim A > \dim W + \dim T - \dim X.$$

The *atypical set* of W (in X) is the union of all atypical components (of W in X), and is denoted $\mathrm{Atyp}(W, X)$, or $\mathrm{Atyp}(W)$ if X is implicit from the context.

Variants of the following conjecture, in different settings, were made independently by Zilber [35], Bombieri-Masser-Zannier [4], and Pink [26].

Conjecture 2.3 (Zilber-Pink for X). Let $W \subset X$. Then $\mathrm{Atyp}(W)$ is a finite union of atypical components; equivalently, there are only finitely many maximal atypical components.

The full Zilber-Pink conjecture is the same statement about an arbitrary mixed Shimura variety (with its special subvarieties), and an algebraic subvariety $W \subset X$. In fact the above is the analogue of the statements in [4, 35] in the general setting considered by Pink, and is notionally stronger than the statement in [26]. For a general discussion of the conjecture see [34].

Definition 2.4. Let $A \subset X$ be a subvariety. We denote by $\langle A \rangle$ the smallest special subvariety containing A (which exists as it is just the intersection of all special subvarities containing A), and define the *defect* of A by

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

Thus $A \subset W$ is atypical if $\delta(A) < \dim X - \dim W$, and W itself is atypical if $\langle W \rangle \neq X$.

Now in Conjecture 2.3 we only look for maximal atypical components, and we do not care if a larger atypical component contains a smaller but more atypical (*i.e.* smaller defect) one. But in fact the conjecture (taken over all special subvarieties of X) implies a formally stronger version (see [14, Proposition 2.4]).

Definition 2.5. A subvariety $W \subset V$ is called *optimal* for V if there is no strictly larger subvariety $W \subset W' \subset V$ with $\delta(W') \le \delta(W)$.

Conjecture 2.6. Let $V \subset X$. Then V has only finitely many optimal subvarieties.

For a particular V and X, finding (or establishing the finiteness of) all optimal subvarieties could be more difficult than finding (or establishing the finiteness of) all maximal atypical subvarieties.

Now (as in [14]) we can repeat the same pattern of definitions with weakly special subvarieties instead of special ones. The smallest weakly special subvariety containing W we denote $\langle W\rangle_{\text{geo}}$, and we define the *geodesic defect* to be

$$\delta_{\text{geo}}(W) = \dim\langle W\rangle_{\text{geo}} - \dim W.$$

A subvariety $W \subset V$ is called *geodesic-optimal* if there is no strictly larger subvariety $W' \subset V$ with $\delta_{\text{geo}}(W') \le \delta_{\text{geo}}(W)$. (This property is termed "cd-maximal" in the multiplicative setting in [27]). The following fact was established for modular, multiplicative and Abelian varieties separately in [14].

Proposition 2.7. *Let $V \subset X_{m,n}$. An optimal subvariety of V is geodesic-optimal.*

Proof. It is easy to adapt the proof of [14, Proposition 4.3] to show that $X_{m,n}$ has the "defect condition", and then the above follows from the formal properties of weakly special and special subvarieties, as in [14, Proposition 4.5]. □

Now we consider

$$V = V_n = \{(x_1, \dots, x_n, t_1, \dots, t_n) : x_i = t_i, i = 1, \dots, n\} \subset X_n = X_{n,n}.$$

We see that if a tuple $(\sigma_1, \dots, \sigma_n)$ of singular moduli satisfies a non-trivial multiplicative relation then the point

$$\Sigma = (\sigma_1, \dots, \sigma_n, \sigma_1, \dots, \sigma_n) \in V$$

lies in the intersection of V with a special subvariety of X of codimension $n + 1$. So such a point is an atypical component of V_n.

3. Mixed Ax-Schanuel

We now take again

$$X = X_{m,n} = Y(1)^m \times \mathbb{G}_{\text{m}}^n, \quad U = U_{m,n} = \mathbb{H}^m \times \mathbb{C}^n, \quad \text{and} \quad \pi : U \to X$$

given by

$$\pi(z_1, \dots, z_m, u_1, \dots, u_n) = \big(j(z_1), \dots, j(z_m), \exp(u_1), \dots, \exp(u_n)\big).$$

Definition 3.1.

1. An *algebraic subvariety* of U will mean a complex-analytically irreducible component of $Y \cap U$ where $Y \subset \mathbb{C}^m \times \mathbb{C}^n$ is an algebraic subvariety;
2. A *weakly special subvariety* of U is an irreducible component of $\pi^{-1}(W)$ where W is a weakly special subvariety of X. Likewise for *special subvariety* of U.

The following result leads to the analogue of the "Weak Complex Ax" (WCA; [14, Conjecture 5.10]) in this mixed modular-multiplicative setting. It is deduced from the same statement in the two extreme special cases: WCA for $Y(1)^n$, which is a consequence of the full modular Ax-Schanuel result established in [24], and WCA for $\mathbb{G}_{\mathrm{m}}^n$, which is a consequence of Ax-Schanuel [1].

Note that we could avoid talking about "algebraic subvarieties of U" by taking Y to be an algebraic subvariety of $\mathbb{C}^m \times \mathbb{C}^n$ and A to be a complex-analytically irreducible component of $Y \cap \pi^{-1}(V)$.

Theorem 3.2. *Let $V \subset X$ and $W \subset U$ be algebraic subvarieties and $A \subset W \cap \pi^{-1}(V)$ a complex-analytically irreducible component. Then*

$$\dim A = \dim V + \dim W - \dim X\,,$$

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

Proof. We suppose that A is not contained in a proper weakly special subvariety of U, and prove the dimension statement. We may suppose that A is Zariski-dense in W and that $\pi(A)$ is Zariski-dense in V.

Let V_0 be the image of V under the projection $X \to \mathbb{G}_{\mathrm{m}}^n$, and W_0 the image of W under the projection $p_0 : U \to \mathbb{C}^n$. Then the image A^* of A under p_0, being connected, is contained in some complex-analytically irreducible component $A_0 \subset W_0 \cap \exp^{-1}(V_0)$. Then A_0 is not contained in a proper weakly special subvariety of $\mathbb{C}^n$, otherwise A would be contained in a proper weakly special subvariety of U. So by Ax-Schanuel ([1]; see also [32]) we have

$$\dim A_0 \le \dim W_0 + \dim V_0 - \dim \mathbb{C}^n.$$

Now we look at fibres in $\mathbb{H}^m$ and $\mathbb{C}^m$. We let $A_u, W_u \subset \mathbb{H}^m$, $V_t \subset \mathbb{C}^m$ be the fibres (of A, W, V respectively) over $u = (u_1, \ldots, u_n) \in A_0$, $u \in W_0$, $t = (t_1, \ldots, t_n) \in V_0$, respectively. Now A_0 must be Zariski-dense in W_0, else A could not be Zariski-dense in W, and similarly $\exp(A_0)$ must be Zariski-dense in V_0.

Since A is irreducible, the image A^* has constant dimension (see [16, V. 3.2, Corollary 2]) equal to the rank $\mathrm{rk}(p_0)$ of $p_0 : A \to A_0$, and $\dim A^* \le \dim A_0$. Further we have [16, V. 3.3] that $\dim A = \mathrm{rk}(p_0) + \lambda(p_0) = \dim A^* + \lambda(p_0)$ where $\lambda(p_0)$ is the generic (*i.e.* minimal) fibre dimension of p_0.

The projection $W \to W_0$ has a generic fibre dimension away from a locus $W' \subset W$ of lower dimension, which does not contain A. So a generic fibre over A^* outside the image of W' is generic for A^* as well as W_0, and likewise for the corresponding fibre over V_0.

For $u \in A^*$, if A_u is not contained in a proper weakly special subvariety of $\mathbb{H}^m$, then by Ax-Schanuel for the j-function [24] we have,

$$\dim A_u \leq \dim W_u + \dim V_u - \dim \mathbb{H}^m.$$

If this holds generically, adding up the two last displays gives us the statement we want.

So we consider what happens when this fails generically. If the A_u were contained in a fixed proper weakly special, than A would be, which we have precluded. So the fibres must belong to a "moving family" of proper weakly specials. As elements of $\mathrm{GL}_2^+(\mathbb{Q})$ can't vary analytically, the only possibility is that some coordinates are constant on the fibres (though not constant on A).

Without loss of generality, we can suppose that these coordinates are $z_1, \ldots, z_k$. For $1 \leq \ell \leq k$, let V_ℓ be the image of V under the projection $X \to \mathbb{C}^\ell \times \mathbb{G}_\mathrm{m}^n$, and W_ℓ the image of W under the projection $p_\ell : U \to \mathbb{H}^\ell \times \mathbb{C}^n$. Then the image of A under p_ℓ, being connected, is contained in some complex-analytically irreducible component $A_\ell \subset W_\ell \cap \pi_{\ell,n}^{-1}(V_\ell)$. Note that this is consistent with the earlier definition of A_0, W_0, V_0.

Now we prove inductively that the dimension inequality holds at "level" ℓ, and once it holds at level k we are done. We assume that, for some $0 \leq h < k$:

(A) A_h is Zariski-dense in W_h and $\pi_h(A_h)$ is Zariski-dense in V_h;
(B) $\dim A_h \leq \dim W_h + \dim V_h - (n + h)$.

We know that these both hold for $h = 0$, and that (A) holds for all h.

Now z_{h+1} is constant on the fibres, so $\dim A_{h+1} = \dim A_h$. To show (B) we need only show that either $\dim W_{h+1} > \dim W_h$ or $\dim V_{h+1} > \dim V_h$.

Suppose that $\dim W_{h+1} = \dim W_h$. This means that, as functions on W, z_{h+1} is algebraic over $z_1, \ldots, z_h, u_1, \ldots, u_n$. But, as W is not contained in a proper weakly special subvariety, z_{h+1} is not constant on W nor does it satisfy any relation $z_{h+1} = gz_i$ where $1 \leq i \leq h$ and $g \in \mathrm{GL}_2^+(\mathbb{Q})$. But then, by the "Ax-Lindemann" result of [22] for the j-function, $j(z_{h+1})$ is algebraically independent of $j(z_1), \ldots, j(z_h), \exp(u_1), \ldots, \exp(u_n)$ as functions on W. Hence by the Zariski density these functions are independent as functions on A_{h+1}, and hence, by the Zariski-density of $\pi_{h+1}(A_{h+1})$ in V_{h+1}, we must have that $\dim V_{h+1} = \dim V_h + 1$. □

From this statement one may deduce, as explained in [23, above 5.7], the analogue of [14, Conjecture 5.10] (for j itself this follows from [24]).

Theorem 3.3. *Let $U' \subset U$ be a weakly special subvariety, and put $X' = \pi(U')$. Let $V \subset X'$ and $W \subset U'$ be subvarieties, and A a complex-analytically irreducible component of $W \cap \pi^{-1}(V)$. Then*

$$\dim A = \dim V + \dim W - \dim X'$$

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

It is shown in [14] that Theorem 3.2 is equivalent by arguments using only the formal properties of the collection of weakly special subvarieties to the following version. We need the following definition from [14].

Definition 3.4. Fix a subvariety $V \subset X$.

1. A *component* with respect to V is a complex analytically irreducible component of $W \cap \pi^{-1}(V)$ for some algebraic subvariety $W \subset U$;
2. If A is a component with respect to V we define its *defect* to be $\partial(A) = \dim \mathrm{Zcl}(A) - \dim A$, where $\mathrm{Zcl}(A)$ denotes the Zariski closure of A;
3. A component A with respect to V is called *optimal* for V if there is no structly larger component B with respect to V with $\partial(B) \leq \partial(A)$;
4. 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.

Proposition 3.5. *Let $V \subset X$. An optimal component with respect to V is geodesic.*

Proof. The same as the proof that "Formulation A" implies "Formulation B" in [14]. (The proof of the reverse implication is also the same as given there.) ☐

4. Proof of Theorem 1.3

We start by recalling some background on trees and lattices associated to $\mathrm{GL}_2^+(\mathbb{Q})$. Let $T_{\mathbb{Q}} = \mathrm{PSL}_2(\mathbb{Z})\backslash\mathrm{PGL}_2^+(\mathbb{Q})$, where we assume their images are distinct. For a prime number p, $T_{\mathbb{Q}}$ maps into $T_p = \mathrm{PSL}_2(\mathbb{Z}_p)\backslash\mathrm{PGL}_2(\mathbb{Q}_p)$, and embeds into the product of the T_p over all p.

Now $T_{\mathbb{Q}}$ may be identified with the space of $\mathbb{Z}$-lattices in $\mathbb{Q}^2$ up to scaling, by sending g to the lattice spanned by $e_1 g, e_2 g$, where $e_1 = (1, 0), e_2 = (0, 1)$. Likewise, T_p may be identified with the space of $\mathbb{Z}_p$-lattices in $\mathbb{Q}_p^2$ up to scale. Moreover, T_p may be given the structure of a connected $(p + 1)$-regular tree by saying that two lattices L, L' are adjacent if one can scale L' to be inside L with index p. There is a natural right action of $\mathrm{PGL}_2(\mathbb{Q}_p)$ on T_p: it acts on $\mathbb{Q}_p^2$ (up to scaling) in the natural way and thus on the lattices in it.

Since T_p is a tree there is a unique shortest path between any two nodes, and any path between those nodes traverses that path.

Our proof will study curves isogenous to the curve E_0 whose j-invariant is 0. These curves have CM by $\mathbb{Z}[\zeta]$, where $\zeta = \exp(2\pi i/3)$. A point $z \in \mathbb{H}$ with $j(z) = 0$ corresponds to the elliptic curve E_0 together with a basis v_1, v_2 for its integral homology $H_1(E_0, \mathbb{Z})$. For any sub-lattice $L \subset H_1(E_0, \mathbb{Q})$ we can define an elliptic curve E_L isogenous to E_0 which only depends on L up to scale. To do this, scale L until it contains $H_1(E_0, \mathbb{Z})$ and the quotient is cyclic. We can identify $Q_L = L/H_1(E_0, \mathbb{Z})$ with a subgroup of the torsion group of E_0 and take the quotient. Define $T'_{\mathbb{Q}}$ to be the space of lattices in $H_1(E_0, \mathbb{Q})$, up to scaling, and correspondingly T'_p the space of $\mathbb{Z}_p$-lattices in $H_1(E_0, \mathbb{Q}_p)$, up to scaling.

Now suppose that E_L is isomorphic to E_0. This implies that the quotient Q_L is the same as that of the kernel of an endomorphism x of E_0. If we identify $H_1(E, \mathbb{Z})$ with $\mathbb{Z}[\zeta]$, then the kernel of multiplication by x is $(x^{-1})/\mathbb{Z}[\zeta]$, where (m) denotes the fractional ideal generated by m. These correspond to elements of the fractional ideal group of $\mathbb{Z}_p[\zeta]$ (providing the endomorphisms giving the kernels) quotiented out by $\mathbb{Q}_p^\times$ (scaling). Explicitly we find the following.

1. If $p \equiv 1 \bmod 3$ then (p) has two disctinct primes above it, whose product is (p). Then $\mathbb{Z}_p[\zeta] = \mathbb{Z}_p \oplus \mathbb{Z}_p$ with ideal group $\mathbb{Z}^2$, which we quotient by the diagonal $\mathbb{Z}$. These nodes give a line in the tree: each such node being adjacent to two other such nodes for which the edges correspond to the two primes over (p);
2. If $p \equiv 2 \bmod 3$ then $\mathbb{Z}_p[\zeta] = \mathbb{Z}_{p^2}$, with ideal group $\mathbb{Z}$ which we quotient by $\mathbb{Z}$. Thus in this case there is just one node coming from curves isomorphic to E_0;
3. If $p = 3$ we get a ramified extension of $\mathbb{Z}_3$, which still has ideal group $\mathbb{Z}$ (generated by powers of the uniformiser) but now we quotient by $2\mathbb{Z}$ since 3 has valuation 2. We thus have two nodes coming from curves isomorphic to E_0, which are adjacent in the tree.

Note that in every case there is at least one node N' of T'_p adjacent to $H_1(E_0, \mathbb{Z})$ such that any lattice L for which the shortest path from $H_1(E_0, \mathbb{Z})$ to L goes through N' is not isomorphic to E_0.

Proposition 4.1. *Let $g_1, \dots, g_k \in \mathrm{GL}_2^+(\mathbb{Q})$ and suppose that the functions $j(g_i z)$ are distinct. Then there exists $z \in \mathbb{H}$ such that $j(g_i z) = 0$ for exactly one i.*

Proof. Suppose first that there exists a prime number p such that the images u_i of the g_i in T'_p are distinct. Without loss of generality we may assume that g_1, g_2 have images u_1, u_2 in T'_p whose distance is at least as large as that between the images of any distinct g_i, g_k. This implies there is a unique node N adjacent to u_1 such that the shortest path from u_1 to any other u_i goes through N. We may further suppose without loss of generality that $g_1 = 1$.

Fixing a basis v_1, v_2 for $H_1(E_0, \mathbb{Z})$ gives a map from T_p to T'_p, sending $\mathbb{Z}^2$ to $H_1(E_0, \mathbb{Z})$. By choosing v_1, v_2 appropriately we can send N to N'. It follows that the z with $j(z) = 0$ corresponding to this choice has $j(g_i z) \neq 0$ for all $i > 1$.

Now we give the proof without the simplifying assumption. While no single p may separate all the g_i, finitely many p do. Let $S = \{g_1, \dots, g_k\}$. Consider the image of S in T_2 and pick two nodes with maximal distance among images of pairs from S. Let u_2 be one of these "extremal" nodes, and let S_2 be the subset of S whose image in T_2 is u_2.

Now consider the image of S_2 in T_3, choose an extremal node u_3 and let S_3 be the subset of S_2 whose image in T_3 is u_3. After finitely many steps we arrive at a set S_p with only a single element. We may assume this element is g_1 and that $g_1 = 1$.

For each prime $q \leq p$ we let N_q be the unique node adjacent to u_q through which all paths from u_q to other images S_r go, where r is the prime preceding q (or $r = 0$ for $p = 2$).

Choose a basis v_1, v_2 of $H_1(E_0, \mathbb{Z})$ such that the induced map from T_q to T'_q takes N_q to N'_q for all $q \le p$. The fact that this is possible amounts to the fact that $\mathrm{SL}_2(\mathbb{Z})$ subjects onto $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$ for every n.

The claim now is that, for each $i > 1$, $j(g_i z) \neq 0$. To see this, let $q < p$ be the largest prime such that $g_i \in S_q$, and $q' \le p$ the next prime after q. The above argument in the tree $T'_{q'}$ shows that $g_i z$ does not represent E_0. This proves the claim and the proposition follows. □

Proof of Theorem 1.3. Theorem 1.3 follows directly from Proposition 4.1 □

5. Arithmetic estimates

The proof of Theorem 1.2, and further results considered in the sequel, use some basic arithmetic estimates which are gathered here. Several of them were used for similar purposes in [13]. The absolute logarithmic Weil height of a non-zero algebraic number α is denoted $h(\alpha)$; the absolute Weil height is $H(\alpha) = \exp h(\alpha)$.

Constants $c_0, c_1, c_2 \ldots$ here and in the sequel are positive and absolute (though not necessarily effective), and have only the indicated dependencies.

Weak Lehmer inequality

A lower bound for the height by any fixed negative power of the degree suffices for our purposes. Loher has proved (see [17]): if $[K : \mathbb{Q}] = d \ge 2$ and $0 \neq \alpha \in K$ is not a root of unity then

$$h(\alpha) \ge \frac{1}{37} d^{-2} (\log d)^{-1}. \tag{5.1}$$

Singular moduli

For a singular modulus σ, we denote by R_σ the associated quadratic order and $D_\sigma = D(R_\sigma)$ its discriminant. Habegger [12, Lemma 1] shows that

$$h(\sigma) \ge c_1 \log |D_\sigma| - c_0, \tag{5.2}$$

based on results of Colmez and Nakkajima-Taguchi.

No singular modulus is a root of unity (we thank Gareth Jones for pointing this out: a singular modulus has a Galois conjugate which is real, but ± 1 are not singular moduli by inspecting the list of rational singular moduli, *e.g.*, in [29, A. 4]). This together with Kronecker's theorem imply, for a non-zero singular modulus σ,

$$h(\sigma) > c_2. \tag{5.3}$$

In the other direction [13, Lemma 4.3], for all $\epsilon > 0$,

$$h(\sigma) \le c_3(\epsilon) |D_\sigma|^\epsilon. \tag{5.4}$$

Finally, we note that if τ is a pre-image of a singular modulus σ in the classical fundamental domain for the $\mathrm{SL}_2(\mathbb{Z})$ action then (see [22, 5.7])

$$H(\tau) \leq 2D_\sigma. \tag{5.5}$$

Class numbers of imaginary quadratic fields

The class number of an imaginary quadratic order R will be denoted $\mathrm{Cl}(R)$. Recall that, for a singular modulus σ, $[\mathbb{Q}(\sigma) : \mathbb{Q}] = \mathrm{Cl}(R_\sigma)$. By Landau-Siegel, for every $\epsilon > 0$,

$$\mathrm{Cl}(R) \geq c_4(\epsilon)|D(R)|^{\frac{1}{2}-\epsilon}. \tag{5.6}$$

In the other direction,

$$\mathrm{Cl}(R) \leq c_5(\epsilon)|D(R)|^{\frac{1}{2}+\epsilon} \tag{5.7}$$

with $c_5(\epsilon)$ explicit (see, *e.g.*, Paulin [20, Proposition 2.2], for a precise statement).

Faltings height of an elliptic curve

Let E be an elliptic curve defined over a number field. Let $h_{\mathrm{F}}(E)$ denote the semi-stable Faltings height of E, and j_E its j-invariant. Then ([31, 2.1]; see also [10])

$$\left| h(j_E) - \frac{1}{12} h_{\mathrm{F}}(E) \right| \leq c_6 \log \max\{2, h(j_E)\} \tag{5.8}$$

with an absolute constant c_6.

Further, if E_1, E_2 are elliptic curves defined over a number field with a cyclic isogney of order N between them (*i.e.* $\Phi_N(j_{E_1}, j_{E_2}) = 0$) then ([28, 2.1.4]; see also [13, proof of Lemma 4.2])

$$|h_{\mathrm{F}}(E_1) - h_{\mathrm{F}}(E_2)| \leq \frac{1}{2} \log N. \tag{5.9}$$

Isogeny estimate

Let K be a number field with $d = \max\{2, [K : \mathbb{Q}]\}$. Let E, E' be elliptic curves defined over K, with $h_{\mathrm{F}}(E)$ and $h_{\mathrm{F}}(E')$ their semi-stable Faltings heights. When E and E' are isogenous, the fundamental results of Masser and Wüstholz [19] give an estimate for the degree of as minimal isogeny between E, E' in terms of $[K : \mathbb{Q}]$ and the height of one of them. Gaudron and Rémond [8] prove the following explicit result improving that of Pellarin [21].

If E, E' are isogenous then there exists an isogeny $E \to E'$ of degree N satisfying

$$N \leq 10^{13} d^2 \max\{h_{\mathrm{F}}(E), \log d, 1\}^2. \tag{5.10}$$

In particular there exists a cyclic isogeny with the same degree bound.

Estimate for the height of a multiplicative dependence

The following result, due to Yu (see [17]), allows us to get control of the height of a multiplicative relation on our singular moduli in terms of their height. It is thus a kind of "multiplicative isogeny estimate".

Let $\alpha_1, \dots, \alpha_n$ be multiplicatively dependent non-zero elements of a number field K of degree $d \geq 2$. Suppose that any proper subset of the α_i is multiplicatively independent. Then there exist rational integers $b_1, \dots, b_n$ with $\alpha_1^{b_1} \dots \alpha_n^{b_n} = 1$ and

$$|b_i| \leq c_7(n) d^n \log d\, h(\alpha_1) \dots h(\alpha_n)/h(\alpha_i), \quad i = 1, \dots, n. \tag{5.11}$$

6. Proof of Theorem 1.2

Fix n. Let $X = X_n = X_{n,n} = Y(1)^n \times \mathbb{G}_{\mathrm{m}}^n$, and let

$$V = V_n = \{(x_1, \dots, x_n, t_1, \dots, t_n) \in X : t_i = x_i, i = 1, \dots, n\}.$$

So $\dim V = \operatorname{codim} V = n$ and a singular-dependent n tuple $(x_1, \dots, x_n)$ gives rise to an atypical point $(x_1, \dots, x_n, x_1, \dots, x_n) \in V$.

Lemma 6.1. *A singular-dependent n-tuple may not be contained in an atypical component of V of positive dimension.*

Proof. A singular-dependent tuple can never be contained in a special subvariety of X defined by two (independent) multiplicative conditions, for between them we could eliminate one coordinate, contradicting the minimality.

Now a special subvariety of the form $M \times \mathbb{G}_{\mathrm{m}}^n$, where M is a special subvariety of $Y(1)^n$, can never intersect V atypically; neither can one of the form $Y(1)^n \times T$ where T is a special subvariety of $\mathbb{G}_{\mathrm{m}}^n$.

Let us consider a special subvariety of the form $M \times T$ where T is defined by one multiplicative condition. The intersection of $M \times T$ with V consists of those n-tuples of M which belong to T. This would typically have dimension $\dim M - 1$, and so to be atypical we must have $M \cap \mathbb{G}_{\mathrm{m}}^n \subset T$. Now Theorem 1.3 implies that M has two identically equal coordinates, but then cannot contain a singular-dependent tuple. □

Proof of Theorem 1.2. If $\sigma = j(\tau)$ is a singular modulus, so that $\tau \in \mathbb{H}$ is quadratic over $\mathbb{Q}$, we define its *complexity* $\Delta(\sigma)$ to be the absolute value of the discriminant of τ *i.e.* $\Delta(\sigma) = |D_\sigma| = |b^2 - 4ac|$ where $ax^2 + bx + c \in \mathbb{Z}[x]$ with $(a, b, c) = 1$ has τ as a root. For a tuple $(\sigma_1, \dots, \sigma_n)$ of singular moduli we define the complexity of σ to be $\Delta(\sigma) = \max(\Delta(\sigma_1), \dots, \Delta(\sigma_n))$.

Now suppose that V contains a point corresponding to a singular-dependent n-tuple of sufficiently large complexity, Δ. By Landau-Siegel (5.6) with $\epsilon = 1/4$, such a tuple has, for sufficiently large (though ineffective) Δ, at least $c_5\Delta^{1/4}$ conjugates over $\mathbb{Q}$. Each is a singular-dependent n-tuple, and they give rise to distinct points in V.

Let F_j be the standard fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{H}$, and $F_{\exp}$ the standard fundamental domain for the action of $2\pi i\mathbb{Z}$ (by translation) on $\mathbb{C}$.

We now consider the sets

$$Y = \left\{(z, u, r, s) \in F_j^n \times F_{\exp}^n \times \mathbb{R}^n \times \mathbb{R} : j(z) = \exp(u), r \cdot u = 2\pi i s\right\},$$

so that $(j(z), \exp(u)) \in V$ for $(z, u, r, s) \in Y$ and

$$Z = \left\{(z, r, s) \in F_j^n \times \mathbb{R}^n \times \mathbb{R} : \exists u, (z, u, r, s) \in Y\right\}.$$

Then Z is a definable set in the o-minimal structure $\mathbb{R}_{\mathrm{an}\,\exp}$.

A singular-dependent n-tuple $\sigma \in V$ has a pre-image

$$\tau = (z_1, \ldots, z_n, u_1, \ldots, u_n) \in F_j^n \times F_{\exp}^n,$$

and this gives rise to a point in Z, where the coordinates in $\mathbb{R}^{n+1}$ register the multiplicative dependence of the tuple, as follows. The F_j coordinates are the z_i, so they are quadratic points, and as recalled in (5.5) their absolute height is bounded by $2\Delta(\sigma_i)$. The point in $\mathbb{R}^{n+1}$ has integer coordinates $(b_1, \ldots, b_n, b)$, not all zero, such that

$$\sum_{i=1}^{n} b_i u_i = 2\pi i b.$$

Now in view of the height estimate (5.4), and degree estimate (5.7) on the $j(z_i)$, (5.11) gives that the b_i in a multiplicative relation among the σ_i may be taken to be bounded in size by $c_8(n)\Delta^n$. With this bound on the $|b_i|$, and since the imaginary parts of the u_i are bounded by $2\pi i$, we get an upper bound on $|b|$. We find that the height of $(z_1, \ldots, z_n, b_1, \ldots, b_n, b)$ is bounded by $c_9(n)\Delta^n$.

In view of the Galois lower bound, a singular-dependent n-tuple of complexity Δ gives rise to at least

$$T^{\frac{1}{4n}} \text{ quadratic points on } Z \text{ with absolute height at most } T = c_{10}(n)\Delta^n.$$

For sufficiently large Δ, the Counting Theorem [25] applied to quadratic points on Z (considered in real coordinates) implies that it contains a semi-algebraic set of positive dimension. This implies (by the arguments used in [13,14]: the corresponding points (z, u) in $\mathbb{H}^n \times \mathbb{C}^n$ cannot be constant on all such semi-algebraic sets) that there is a complex algebraic $Y \subset U$ which intersects Z in a positive-dimensional component A which is atypical in dimension and contains singular-dependent n-tuples.

By the mixed Ax-Schanuel of Section 3 this implies that there is a positive-dimensional weakly special subvariety W containing Y containing a component B with $A \subset B$ and $\partial(B) \leq \partial(A)$. Moreover, it contains the special subvarieties that contain (some of) the singular-dependent points, so W is a special subvariety of positive dimension containing singular-dependent points of V, which we have seen is impossible.

So Δ is bounded, giving the finiteness. $\square$

Example 6.2. An example of a singular-dependent 5-tuple is (see [29, A. 4]):

$$\big(-2^{15}3^3 5^3 11^3, \quad -2^{15}, \quad 2^3 3^3 11^3, \quad 2^6 3^3, \quad 2^{15}3^1 5^3\big).$$

One also has a 3-tuple (-2^{15}, $-2^{15}3^3$, $2^6 3^3$) and 4-tuple ($2^4 3^3 5^3$, $-2^{15}3^1 5^3$,$-3^3 5^3$, $2^6 5^3$).

7. On the atypical set of V_n

The atypical set of V_n is the union of its proper optimal components (V_n itself is always optimal but never atypical). Since optimal components are geodesic-optimal Proposition 2.7, we will investigate the possibilities for these.

We observe that any geodesic-optimal components which dominate every coordinate can only come from an optimal strongly special subvariety. The finiteness of these, even if we cannot identify them, is guaranteed by o-minimality.

Definition 7.1. Complex numbers x, y will be called *Hecke equivalent* if $\Phi_N(x, y) = 0$ for some $N \geq 1$. *I.e.*, if the elliptic curves with j-invariants x and y are isogenous.

7.1. Geodesic-optimal components of dimension n

As already observed, V_n is not atypical since it dominates both $Y(1)^n$ and $\mathbb{G}_{\mathrm{m}}^n$. In other words, the defect of V_n is equal to its codimension.

7.2. Geodesic-optimal components of dimension $n - 1$

Let $T \subset X$ be a geodesic subvariety of co-dimension 2. Can $T \cap V$ have dimension $n - 1$? There are two equations defining T, each being one of four possible types: a single modular relation, a constant modular coordinate, a single multiplicative relation, a constant multiplicative coordinate.

Now if both equations are of modular (respectively multiplicative) type we never get an atypical component, because V dominates $Y(1)^n$ (respectively $\mathbb{G}_{\mathrm{m}}^n$). The same is true for any T which is defined purely by modular (respectively multiplicative) relations.

So we consider T defined by one condition of each type. Let us call T_1 the projection of T to the $Y(1)^n$ factor, which is a geodesic subvariety of codimension 1, and T_2 its projection to $\mathbb{G}_{\mathrm{m}}^n$. We get an atypical component of dimension $n - 1$ if either $T_1 \cap \mathbb{G}_{\mathrm{m}}^n$ is contained in T_2, or if T_2 is contained in T_1 (*i.e.* when both are considered in the same copy of $(\mathbb{C}^\times)^n$).

If the modular condition is a modular relation (rather than a constant coordinate) then the first is excluded by Theorem 1.3, unless it is of the form $x_i = x_j$. If the multiplicative relation is not a fixed coordinate, the other inclusion is also impossible unless it is of the form $t_i = t_j$.

So we are reduced to considering constant coordinate conditions on both sides. This obviously leads to a component of dimension $n - 1$ if the conditions coincide:

$x_i = \xi = t_i$. However such a component can only be atypical (*i.e.* arise from the intersection of V_n with a special subvariety of codimension (at most) 2 if ξ is both a singular modulus and a root of unity. But this never occurs, as remarked in Section 5.

This establishes ZP for V_1, which is the curve defined by $x_1 = t_1$ in $\mathbb{C} \times \mathbb{C}^\times$. And it shows that V_2 has no atypical subvarieties of positive dimension apart from the "diagonal" $x_1 = x_2$.

Proposition 7.2. *ZP holds for* V_2.

Proof. In view of the fact that the only atypical component of positive dimension is the "diagonal", which has defect zero, we are reduced to showing that V_2 has only finitely many optimal points, *i.e.* points which are atypical but not contained in the "diagonal". A point $(x_1, x_2, x_1, x_2) \in V_2$ is atypical if it lies on a special subvariety of codimension 3. There are then two cases: we have two independent modular conditions and one multiplicative, or two multiplicative and one modular relation.

The former case is exactly the question of singular-dependent 2 tuples, whose finiteness we have already established. The latter leads to the question of two (unequal) roots of unity which satisfy a modular relation. This is established in the following proposition, by a similar argument to that used in (5.2); and with this the proof is complete. □

We may observe that the optimal points of V_2 satisfy 3 special relations (never 4), so have "defect" 1.

Definition 7.3. A pair of distinct roots of unity is called a *modular pair* if they satisfy a modular relation.

Proposition 7.4. *There exist only finitely many modular pairs.*

Proof. Let (ζ_1, ζ_2) be such a point, where the order of ζ_i is M_i and $\Phi_L(\zeta_1, \zeta_2) = 0$. The point is that the order of the root of unity, and their bounded height, leads to a bound on the degree of the modular relation. Specifically, by (5.8), the semistable Faltings height of the corresponding elliptic curves E_1, E_2 with j-invariants ζ_1, ζ_2 are bounded, and so by the isogeny estimate (5.10) there is a modular relation $\Phi_N(\zeta_1, \zeta_2) = 0$ with $N \leq c_{11} \max\{M_1, M_2\}^5$. Thus such a point leads to a rational point on a suitable definable set whose height is bounded by a polynomial in the orders of the two roots, and if it is of sufficiently large complexity it forces the existence of a higher dimensional atypical intersection containing such points. But the only atypical set of dimension 1 is given by $x_1 = x_2, t_1 = t_2$. □

As modular relations always subsist between two numbers, there is no notion of "modular-multiplicative n-tuples" analogous to singular-dependent tuples. However, an immediate consequence of the above is that, for any n, there exists only finitely many n-tuples of distinct roots of unity which are pairwise Hecke equivalent (and none for sufficiently large n).

7.3. Geodesic-optimal components of dimension $n-2$

These arise from intersecting V_n with a geodesic subvariety T of codimension (at least) 3. We must have at least 1 relation of each type, and if they are all of "non-constant" type (no fixed coordinates) then we get finiteness by o-minimality.

If there is one constant condition, this immediately gives a second such condition of the other type, and then any additional non-constant condition (*i.e.* not forcing any further constant coordinates) will give a component of dimension $n-2$. However, no such component can be atypical.

Consider the case of 3 constant conditions. First the case of two fixed modular coordinates. This will give rise to an atypical intersection if the two fixed values are multiplicatively related. Next the case of two fixed multiplicative coordinates. This will give rise to an atypical component if the two fixed values are Hecke equivalent. The finiteness of such components follows from ZP for V_2, and they all have defect 2. Thus:

Proposition 7.5. *For $n \geq 1$, V_n has only finitely many maximal atypical components of dimension $n-2$.*

But for $n = 3$ we can in fact exclude "strongly atypical" altogether. Such a component has one of two shapes.

1. Two modular relations and one multiplicative relation. This would be atypical if the resulting modular curve satisfied the multiplicative relation, but this is impossible by Theorem 1.3;
2. Two multiplicative relations and one modular relation. This gives a "multiplicative curve", which can be parameterised as $(\zeta_1 t^{a_1}, \zeta_2 t^{a_2}, \zeta_3 t^{a_3})$, where ζ_i are roots of unity and a_i integers. As the $\Phi_N, N \geq 2$ are symmetric, two coordinates cannot satisfy a modular equation unless $a_i = a_j$ (so that $N = 1$ and $\Phi_1 = X - Y$) and $\zeta_i = \zeta_j$.

Proposition 7.6. *The positive dimensional atypical components of V_3 and their defects may be described as follows:*

1. *The intersection of V_3 with $x_i = x_j, i \neq j$ is a copy of V_2 contained in X_2 (hence of defect 2) and has some atypical points in it, which have defect 1. It contains also the subvariety with $x_1 = x_2 = x_3$, which has defect 0;*
2. *A singular-dependent 2-tuple $\sigma = (\sigma_1, \sigma_2)$ gives rise to an atypical component A_σ of dimension 1 and defect 2. (There may exist singular moduli which belong to two distinct such pairs σ, σ'. Then we get a point $(A_\sigma \cap A_{\sigma'})$ of defect 1);*
3. *A modular pair $\zeta = (\zeta_1, \zeta_2)$ gives rise to an atypical component B_ζ of dimension 1 and defect 2. (There may exist roots of unity belonging to two distinct modular pairs ζ, ζ'. Then we get a point $(B_\zeta \cap B_{\zeta'})$ of defect 1.)*

In particular, there are no positive dimensional "strongly atypical" components (i.e., *with no constant coordinates*).

Thus ZP for V_3 depends on the finiteness of its atypical points off all the above positive dimensional atypical components. This leads to some Diophantine questions which would establish ZP for V_3, which we study in the next section.

Remark 7.7. Note that X contains families of weakly special subvarieties which intersect V_n atypically, namely those defined be relations the form $x_i = x_j$ (and $t_i = t_j$) or $x_k = t_k = c_k \in \mathbb{C}^\times$ for various choices of $(i, j), i \neq j, k$. If m such conditions are imposed, the resulting weakly special subvariety has dimension $2n - 2m$ and intersects V_n in a component of dimension $n - m$, so has geodesic defect $n - m$.

Conjecture 7.8. The atypical geodesic components described in Remark 7.7 give all geodesic optimal subvarieties of V_n for any n; in particular, apart from components defined by "diagonal" equations $x_i = x_j$ there are no "strongly optimal" geodesic optimal components (*i.e.* with no constant coordinates).

8. Optimal points in V_3

The optimal points in $(x_1, x_2, x_3, x_1, x_2, x_3) \in V_3$ fall into two classes. Those which are atypical in satisfying at least 4 special conditions, but are not contained in atypical component of higher dimension. And those which are "more atypical", satisfying 5 special conditions (it is not possible to have 6: only a triple of singular moduli which were also roots of unity could achieve this), though lying in an atypical set of larger dimension but larger defect. Those lying on diagonals $x_i = x_j, i \neq j$ are easy to describe, we consider here those that do not.

Let us first consider points satisfying 5 special conditions. These also fall into two types: 3 modular, 2 multiplicative, or the other way around. If there are 3 modular conditions then each x_i is a singular modulus. The two multiplicative conditions mean either than one x_j is torsion, and the other two multiplicatively related, or the three are pairwise multiplicatively related. The former is impossible. Now only finitely many pairs of singular moduli have a multiplicative relation, so x_1, x_2 comes from a finite set, and x_3 comes also from a finite set. If there are three multiplicative relations then each x_i is torsion. Only finitely many pairs of (distinct) roots of unity satisfy isogenies, and we get finiteness (there are no "Hecke equivalences" involving three points!). All these points have defect 1.

Now we consider points $(x_1, x_2, x_3, x_1, x_2, x_3) \in V_3$, away from positive dimensional atypical subvarieties, satisfying 4 special conditions. The "generic" situation involves no singular moduli or roots of unity.

Problem 8.1. Prove that there exist only finitely many triples x_1, x_2, x_3 of distinct non-zero algebraic numbers, which are not roots of unity and not singular moduli, such that they are pairwise Hecke equivalent, and also pairwise multiplicatively dependent.

The various arithmetic estimates seem insufficient to get a lower degree bound in terms of the "complexity": the degrees of the two isogenies and the heights of the two multiplicative relations. This seems to be a problem of a similar nature to that encountered in [13] dealing with curves which are not "asymmetric" (see [13, Section 1]).

There are three "non-generic" variations of which we can resolve two. The multiplicative relations may take the form that one coordinate is a root of unity, the other two coordinates being multiplicatively dependent. Similarly, the modular relations may take the form that one coordinate is singular, the other two Hecke equivalent. Or both. Note that if two coordinates are singular the point is not optimal, but lies on one of the atypical components in Propositions 7.6(3); if two coordinates are roots of unity, the point is on a component as in Proposition 7.6(2).

We consider the non-generic multiplicative condition first. Up to permutations we may assume the singular coordinate is x_1

Proposition 8.2. *There exist only finitely many triples x_1, x_2, x_3 of distinct non-zero algebraic numbers such that:*

1. *x_3 is a root of unity, x_1, x_2 are multiplicatively dependent but not roots of unity;*
2. *The three points are pairwise Hecke equivalent, but are not singular moduli.*

Proof. Define the complexity Δ of such a triple to be the maximum of: the order M of the root of unity x_3 and the minimum degrees of isogenies N_1, N_2 between x_3 and x_1, x_2, respectively. By (5.8), the stable Faltings height of an elliptic curves whose j-invariant is a root of unity is absolutely bounded. Now by (5.9), $h(x_j) \ll (1 + \log \max\{N_j\})$, $j = 1, 2$, so by (5.10) the degrees $d_j = [\mathbb{Q}(x_3, x_j) : \mathbb{Q}] \gg N_j^{1/5}$. By (5.11) and (5.1) (to get a lower bound for $h(x_i)$) the height of a multiplicative relation between x_1, x_2 is bounded by some $c_{12}\Delta^{c_{13}}$. And $[\mathbb{Q}(x_3) : \mathbb{Q}] = \phi(M) \gg_\epsilon M^{1-\epsilon}$, where ϕ is the Euler ϕ-function. We may take $\epsilon = 1/2$ say.

Thus, a triple of complexity Δ gives rise to "many" (*i.e.* at least $c_{14}\Delta^{c_{15}}$) quadratic points on a certain definable set, and so all but finitely many such points lie on atypical components of positive dimension.

But no such triples lie on positive dimensional atypical components: by Proposition 7.6, such components have either two singular coordinates or two modular coordinates, so the conditions on our triples would then force all x_i to be singular, which is impossible (as then x_3 cannot be torsion) or all torsion, which leads to the same impossible requirement for x_1. □

Symmetrically, we have the case where the modular relations are of the non-generic form. We seem unable to establish finiteness here, so we pose it as a problem.

Problem 8.3. Prove that there exist only finitely many triples x_1, x_2, x_3 of distinct non-zero algebraic numbers such that x_1 is singular, x_2, x_3 are Hecke equivalent, and the three are pairwise multiplicatively dependent.

Finally, we have the following.

Proposition 8.4. *There exist only finitely many triples x_1, x_2, x_3 of distinct non-zero algebraic numbers such that*

1. x_1 *is a singular modulus,* x_2, x_3 *are Hecke equivalent but are not singular moduli;*
2. x_3 *is a root of unity,* x_1, x_2 *are multiplicatively dependent but not roots of unity.*

Proof. Let D be the discriminant of x_1 (see Section 5), and M the (minimal) order of x_3. Take N minimal with $\Phi_N(x_2, x_3) = 0$, and B minimal for a non-trivial multiplicative relation $x_1^{b_1} x_2^{b_2} = 1$ with $B = \max\{b_1, b_2\}$. Set $\Delta = \max\{|D|, M, N\}$ to be the complexity of the tuple (x_1, x_2, x_3). Set $d = [\mathbb{Q}(x_1, x_2, x_3) : \mathbb{Q}]$.

Let E_ξ be the elliptic curve with j-invariant ξ. As in the proof of Proposition 8.2, $h_{\mathrm{F}}(E_{x_3})$ is bounded by some absolute c_{16}. Then, by the isogeny estimates (5.10), we have $N \leq c_{17}([\mathbb{Q}(x_2, x_3) : \mathbb{Q}])^5$. Also $M^{1-\epsilon} \ll_\epsilon \phi(M) = [\mathbb{Q}(x_3) : \mathbb{Q}]$, and $|D| \ll [\mathbb{Q}(x_1) : \mathbb{Q}]^4$ by (5.6).

Arguing as in [13], the height inequalities (5.8), (5.9) imply that $h(x_2)$ is bounded above by $c_{18}(1+\log N)$. By the Weak Lehmer estimate (5.1) it is bounded below by $c_{19} d^{-3}$. Corresponding estimates for $h(x_1)$ are provided by (5.4) and (5.3). Therefore (5.11) ensures that

$$B \leq c_{20} d^3 D.$$

The rest of the proof is the same as the proof of Proposition 8.2. □

Thus Problems 8.1 and 8.3 imply (and are implied by) ZP for V_3. If one takes two complex numbers and three conditions, then either two "modular" conditions or two "multiplicative" special conditions will force the points to be special, and one can prove finiteness. However one can consider two complex numbers satisfying a special condition of each of three (or more) different types.

Let S be a Shimura curve corresponding to a quaternion algebra over $\mathbb{Q}$ (see, *e.g.*, Elkies [6]). There is a notion of *Hecke orbit* of a point on S (see, *e.g.*, [5]), an equivalence class of points under a certain equivalence relation. This relation is the existence of a "cyclic N-isogeny" between the corresponding parameterised objects for some N; see [6, Section 2.3, page 12]. If S has genus zero, there is an analogue $j_S : \mathbb{H} \to \mathbb{P}^1$ of the j-function (see [7,30]) which generates the function field of S, and we may speak of points in $\mathbb{C}$ being "Hecke equivalent (for S)" if they are in the same Hecke orbit.

Problems 8.5. Prove that there are only finitely many pairs of distinct non-zero algebraic numbers x_1, x_2 in each situation.

1. x_1, x_2 are Hecke equivalent (in the sense of 7.1), and multiplicatively dependent, and are also Hecke equivalent for some other Shimura curve;
2. x_1, x_2 are Hecke equivalent, and multiplicatively dependent, and the points with these x-coordinates are dependent in some specific elliptic curve;
3. As in the previous problems, but with more or different conditions: say the points are Hecke equivalent/dependent for 10 pairwise incommensurable Shimura curves.

Finally we state a "weakly bounded height conjecture" on the height of "just likely" intersections of mixed multiplicative-modular type under which Problems 8.1 and 8.3 are affirmed.

Definition 8.6. A *modular-dependent pair* is a point $(x, y) \in (\mathbb{C}^\times)^2$ such that there exists integers N, a, b, ℓ with $N \geq 2$, $\ell \geq 1$ and $\gcd(a, b) = 1$ such that

$$\Phi_N(x, y) = 0, \quad (x^a y^b)^\ell = 1.$$

The *complexity* $\Delta(x, y)$ of such a pair is the minimum of $\max(N, |a|, |b|, \ell)$ over all N, a, b, ℓ for which the above equations hold for x, y.

Conjecture 8.7. For $\epsilon > 0$ we have $h(x), h(y) \leq c_\epsilon \Delta(x, y)^\epsilon$ for all modular-dependent pairs (x, y).

Proposition 8.8. *Assume Conjecture* 8.7. *Then finiteness holds in Problems* 8.1 *and* 8.3.

Proof. Let (x, y) be a modular-dependent pair with complexity $\Delta = \Delta(x, y) = \max(|a|, |b|, \ell)$ for suitable a, b, ℓ. Constants denoted c are absolute but may vary at each occurrence.

Let E_x, E_y be elliptic curves with j-invariants x, y and semistable Faltings heights $h_{\mathrm{F}}(x) = h_{\mathrm{F}}(E_x)$ and $h_{\mathrm{F}}(y) = h_{\mathrm{F}}(E_y)$ respectively. Then E_x, E_y may both be defined over $\mathbb{Q}(x, y)$, and we set $d = [\mathbb{Q}(x, y) : \mathbb{Q}]$.

By the isogeny estimate (5.10), $N \leq cd^2 \max\{h_{\mathrm{F}}(x), \log d, 1\}^2$. Now $h_{\mathrm{F}}(x)$ and $h(x)$ differ by at most $c \log \max(2, h(x))$. So

$$N \leq cd^2 \max(1, \log d)^2 \big(1 + h(x) + c \log \max(2, h(x))\big)^2.$$

We have $d^2 \max(1, \log d)^2 \leq d^4$, and under Conjecture 8.7 (with $\epsilon = 1/20$ say) we have

$$N \leq cd^4 \Delta^{1/10}.$$

For the purposes of Proposition 8.1 and 8.3 we may assume that neither x nor y is a root of unity. By a Weak Lehmer inequality (5.1) we have $h(x) \geq cd^{-3}, \quad h(y) \geq cd^{-3}$. Since neither x, y is a root of unity, we find (5.11) that there exists a nontrivial multiplicative relation $x^\alpha y^\beta = 1$ with

$$|\alpha| \leq cd^3 h(y) \leq cd^3 \Delta^{1/10}, \quad |\beta| \leq cd^3 h(x) \leq cd^3 \Delta^{1/10}.$$

Again since x, y are not roots of unity, we have that (α, β) is a multiple of $(\ell a, \ell b)$. So we find that $|a|, |b|, c\ell \leq cd^3 \Delta^{1/10}$. Now $\Delta = \max(N, |a|, |b|, |\ell|)$ and so combining the various inequalities we find

$$\Delta \leq cd^7.$$

Now points x_1, x_2, x_3 as in Problem 8.1 give rise to rational points on some suitable definable set of height at most $\max\big(\Delta(x_1, x_2), \Delta(x_2, x_3), \Delta(x_1, x_3)\big)$. This lower estimate for the degree is then suitable to complete a finiteness proof for isolated points of this form by point-counting and o-minimality as in the proofs of Theorem 1.2, Propositions 8.2, and 8.4. The argument for Problem 8.3 is similar. □

9. On ZP for V_n

The referee asked us whether there is a natural generalization of the height-theoretic Conjecture 8.7 which would imply ZP for $V_n, n \geq 3$. We thank the referee for raising this question, to which we offer an affirmative answer here. As this Conjecture 9.5 is rather more speculative than the very special case in Conjecture 8.7 we have preferred to keep this section separate.

We continue to let $X_n = Y(1)^n \times \mathbb{G}_{\mathrm{m}}^n$ and $U_n = \mathbb{H}^n \times \mathbb{C}^n$, and F a standard fundamental domain for the action on U_n by $\mathrm{SL}_2(\mathbb{Z})^n \times \mathbb{Z}^n$ where $m \in \mathbb{Z}$ acts on $\mathbb{C}$ be translation by $2\pi i m$. Constants $c, c(n), c(n, \epsilon), \ldots$ depend only on the indicated quantities, but may differ at each occurence.

First, we will assume Conjecture 7.8. This seems to be necessary, in view of the following. We have seen that an optimal component with respect to $V_n \subset X_n$ is geodesic optimal. Now weakly special subvarieties of U_n are contained in larger definable families of "Mobius varieties" which are defined by some finite number of relations of the form $z_i = g_{ij} z_j, g_{ij} \in \mathrm{SL}_2(\mathbb{R})$ or of the form $z_k = c_k \in \mathbb{C}$ on the $\mathbb{H}^n$ variables and of the form $\sigma_{j=1}^n r_{ij} u_j = 0$ with $(r_1, \ldots, r_n) \in \mathbb{R}^n$ on the $\mathbb{C}^n$ variables (see Mobius varieties in [14, Section 6.2] and "linear varieties" in [22, 10.1]). Then the set of Mobius varieties which intersect $Z = \pi^{-1}(V_n) \cap F$ optimally among Mobius varieties gives the full set of weakly special varieties intersecting Z optimally. By o-minimality, the set of relations among non-constant coordinates is then finite, since the corresponding coefficients r_{ij} and group elements g_{ij} must in fact be rational (see [14, Proposition 6.6; 22, 10.2]).

We are thus led to consider, for example, the intersections of a fixed strongly special modular special subvariety M with a family of translates $\{aT : a \in A\}$ of a subtorus T, *i.e.* a family of weakly special multiplicative weakly special subvarieties. Here A can be taken to be a copy of some $(\mathbb{C}^\times)^m$. Since optimal components are geodesic-optimal, these components will correspond to those $a \in A$ for which $M \cap aT$ has atypical dimension, which give some subvarieties $A_i \subset A$. However, if the component is optimal, the corresponding special subvariety will in general be larger, and we will be led to consider atypical points in $A_i \subset (\mathbb{C}^\times)^m$, *i.e.* to some cases of ZP for the multiplicative group, which we do not know how to handle at present.

We will say that a point $C = (x_1, \ldots, x_n) \in (\mathbb{C}^\times)^n$ satisfies h special relations if the smallest special subvariety of X_n containing $(x_1, \ldots, x_n, x_1, \ldots, x_n)$ has codimension h. On the modular side, the relation of being in the same Hecke orbit divides the non-special coordinates into k equivalence classes. Such equivalence classes of non-special points we call *cliques*. Then we see that if C satisfies h special relations we have

$$h = n + m - k,$$

where m is the number of independent multiplicative relations satisfied by C, and k is the number of cliques. We set

$$\partial(C) = n - h = k - m.$$

Definition 9.1. A tuple $C = (c_1, \dots, c_n) \in (\mathbb{C}^\times)^n$ with pairwise distinct coordinates is called *n-optimal* if no proper subtuple C' has $\partial(C') \leq \partial(C)$. *I.e.* removing any ℓ points from C loses at least $\ell + 1$ special relations.

Proposition 9.2. *Assuming Conjecture* 7.8, *ZP for all V_n is equivalent to the statement that, for all n, there are only finitely many n-optimal points.*

Proof. Assuming Conjecture 7.8, all optimal components are, up to permutations of coordinates, of the form $\{(W, W) \in V_n\}$ where $W \subset \mathbb{C}^n$ is of the form

$$\{(c_1, c_2, \dots, c_\ell, x_{\ell+1}, \dots, x_n) : x_i \in \mathbb{C}^*, x_i = x_j : (i, j) \in I\}$$

for some set I of pairs (i, j) with $\max(i, \ell) < j$ from $\{1, \dots, n\}$, where $c_1, \dots, c_\ell$ are distinct non-zero complex numbers.

If the tuple $C = (c_1, \dots, c_\ell)$ satisfies h special relations, we have

$$\dim W = n - \ell - |I|, \quad \dim\langle W\rangle = 2n - 2|I| - h,$$

whence

$$\delta(W) = 2n - 2|I| - h - (n - \ell - |I|) = n + (\ell - h) - |I| = n + \partial(C) - |I|.$$

Therefore, the component W is optimal just if C is ℓ-optimal. □

Suppose C is an n-optimal tuple, with (C, C) contained in some smallest special subvariety $T \subset X_n$. Then the component of $T \cap V_n$ containing (C, C) must be just the point $\{(C, C)\}$. Otherwise, the component is clearly not optimal. Thus, an n-optimal tuple is a tuple of algebraic numbers, and the degree $d(C) = [\mathbb{Q}(c_1, \dots, c_n) : \mathbb{Q}]$ is bounded in terms of the degrees of the equations defining T.

We now frame a "weakly bounded height conjecture" for certain "just likely" intersections that seems plausible and is sufficient to establish this finiteness (assuming Conjecture 7.8). Of course one only needs the conjecture to hold for optimal points, which must in fact be "unlikely".

Consider a point $C = (c_1, \dots, c_n) \in (\mathbb{C}^\times)^n$ with c_i distinct, together with a $\mathbb{Z}$-module Γ of exponents of multiplicative relations on C. That is, Γ is a $\mathbb{Z}$-submodule of the *relation group*

$$\Gamma(C) = \{(a_1, \dots, a_n) \in \mathbb{Z}^n : c_1^{a_1} \cdots c_n^{a_n} = 1\}.$$

Suppose C has k cliques and $\mathrm{rank}(\Gamma) = m$. Removing some points from C yields a tuple C', and it inherits a submodule Γ' of relations from Γ which are trivial on the points removed (*i.e.* Γ' is the submodule of exponent vectors for which the coordinates corresponding to $C - C'$ are zero). We call Γ' the *induced relations*.

Definition 9.3. A pair (C, Γ) consisting of a tuple $C \in (\mathbb{C}^\times)^n$ with pairwise distinct coordinates having k cliques and a $\mathbb{Z}$-module of $\Gamma \subset \mathbb{Z}^n$ of exponent vectors of multiplicative relations on C is called *grounded* if, for any subtuple C' formed by removing any κ cliques, where $0 < \kappa < k$, together with any number of special points, the induced relation module Γ' satisfies $\operatorname{rank}(\Gamma') < \operatorname{rank}(\Gamma) - \kappa$. *I.e.* losing κ cliques loses at least $\kappa + 1$ multiplicative relations.

Note that a grounded tuple can never contain a singleton clique, for omitting such a clique will lead to the loss of at most one multiplicative relation.

Definition 9.4. We define the *height* of a tuple $C = (c_1, \dots, c_n)$ to be

$$h(C) = \max(h(c_1), \dots, h(c_n)).$$

The *modular complexity* of a tuple is

$$\Delta_{\mathrm{mod}}(C) = \max\{N\}$$

over N such that there exists c_i, c_j (allowing $i = j$) with $\Phi_N(c_i, c_j) = 0$, and $N \geq 2$ minimal for this pair i, j. We define the *complexity* of a $\mathbb{Z}$-submodule of $\mathbb{Z}^n$ to be

$$\Delta(\Gamma) = \min\{T\}$$

over T such that there is a basis of Γ consisting of vectors with all entries of absolute value bounded by T. The *complexity* of a pair (C, Γ) is

$$\Delta(C, \Gamma) = \max\{\Delta_{\mathrm{mod}}(C), \Delta(\Gamma)\}.$$

Finally, the *complexity* of C is

$$\Delta(C) = \Delta(C, \Gamma(C)).$$

Conjecture 9.5. Let (C, Γ) be grounded, where $C \in (\mathbb{C}^\times)^n$, and suppose that $\operatorname{rank}(\Gamma)$ equals the number k of cliques of C. Then

$$h(C) \leq c(n, \epsilon)\Delta(C, \Gamma)^\epsilon$$

for any $\epsilon > 0$.

Note that such C is in the intersection of V_n with a special subvariety of dimension $n + k - \operatorname{rank}(\Gamma) = n$, hence is a "just likely" intersection, though this may not be the smallest special subvariety containing C.

It seems that one cannot hope to have a suitable weakly bounded height conjecture for tuples which are not grounded. For example, if one has a clique C' satisfying some relations Γ', then imposing just one additional relation $\gamma \in \mathbb{Z}^n$ on $C = (C', C'')$ for an additional clique C'' would allow the height of C'' to be roughly $\Delta(\mathbb{Z}\gamma)h(C')$. An interesting question seems to be whether one should expect in fact an upper bound of the form $\ll_n (\log \Delta(C, \Gamma))^{c(n)}$ in Conjecture 9.5.

We now gather some further arithmetic estimates. Various forms of the following result, which we do not need in the sharpest forms, appear in the literature; see [17]. The following is extracted from [18].

Proposition 9.6. *For an n-tuple C of degree $d = d(C)$ and height $h = h(C)$, there is a basis of $\Gamma(C)$ consisting of vectors of integers of size at most*

$$c(n)d^n \log(d+2)^{3n} \max(h, 1)^n.$$

Proof. This is a weakened form of the bound in [18, page 253] together with the estimates for the quantities there established on page 254. ☐

Lemma 9.7. *Suppose $C \in (\mathbb{C}^\times)^n$ is n-optimal. Then $(C, \Gamma(C))$ is grounded. Moreover, if C has k cliques, then there is a submodule Γ of relations on C with* $\operatorname{rank}(\Gamma) = k$ *such that (C, Γ) is grounded and $\Delta(\Gamma) \leq c(n)\Delta(\Gamma(C))$.*

Proof. For the first assertion, if we remove κ cliques (and some special points) from C to form C' and lose only κ multiplicative relations then $\partial(C') \leq \partial(C)$, and C was not n-optimal. So n-optimal is stronger than grounded.

For the second assertion, we show how to find a suitable submodule Γ of $\Gamma(C)$ of rank equal to k, the number of cliques.

We first show that there is a vector $v_1 \in \Gamma(C)$ with a height bound as in the assertion of the lemma which "involves" all special points and cliques, *i.e.* where the exponent is non-zero on every coordinate i where c_i is special, and for some coordinate in every clique.

Let B be a basis of $\Gamma(C)$ consisting of vectors of integers of size at most $\Delta(\Gamma(C))$. Since $(C, \Gamma(C))$ is n-optimal, such a vector w_j exists for each individual special coordinate, and for each individual clique; say there are J such vectors. Moreover, we can assume that each $w_j \in B$. We consider vectors of the form

$$w = \sum_{j=1}^{J} a_j w_j, \quad a_j \in \mathbb{Q}.$$

For each special coordinate or clique, the condition that w vanishes on that coordinate or clique gives a proper subspace of $\mathbb{Q}^n$. It therefore contains at most $c(n)T^{n-1}$ integer points in the box $[-T, T]^n$. We must avoid $J \leq n$ such subspaces, so $T = c(n)$ suffices.

We now construct $v_2 \in \Gamma(C)$ such that, for every clique, v_2 does not vanish modulo v_1. For each clique individually the existence of such a vector w_j is assured since $(C, \Gamma(C))$ is grounded, and so we can take $w_j \in B$. A similar box argument produces v_2 (the number of subspaces to avoid is now at most the number of pairs of cliques), and we continue to produce $v_3, \ldots, v_k$, where v_k does not vanish modulo $\mathbb{Z}[v_1, \ldots, v_{k-1}]$ on any choice of $k-1$ cliques. ☐

Proposition 9.8. *Let σ be a singular modulus of discriminant D_σ. Then there exists $N \geq 2$ with $\Phi_N(\sigma, \sigma) = 0$ satisfying $N \leq |D_\sigma|$. Conversely, if $\Phi_N(\sigma, \sigma) = 0$ where $N \geq 2$ then $|D_\sigma| \leq cN^{20}$, with an explicit c.*

Proof. Let τ be a preimage of σ in the standard fundamental domain. Then τ satisfies a minimal quadratic equation over $\mathbb{Z}$ of the form $A\tau^2 + B\tau + C = 0$ which is reduced, meaning $|B| \le A \le C$ and $B \ge 0$ if $A = |B|$ or $A = C$. Thus $4AC = B^2 - D_\sigma \le AC - D_\sigma$ whence $3AC \le |D_\sigma|$. Now $g\tau = \tau$ for $g = \left(\begin{smallmatrix} -B & -C \\ A & 0 \end{smallmatrix}\right)$, which is primitive of determinant $N = AC \le |D_\sigma|$.

In the other direction, suppose $\Phi_N(\sigma, \sigma) = 0$. This means that $g\tau = h\tau$ for a matrix g of the form $g = \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)$ with $0 < a, 0 \le b < d, ad = N$ (see [15, 5.1, page 52]), and $h \in \mathrm{SL}_2(\mathbb{Z})$. Now $|\mathrm{Re}(g\tau)| \le 2N$ and $|\mathrm{Im}(\tau)^{-1}| \le 2N$ so the matrix h has entries at most $c(2N)^9$ by [13, Lemma 5.1]. Thus τ is fixed by $h^{-1}g$, an integer matrix with entries bounded by $c(2N)^{10}$. This gives an integral quadratic polynomial satisfied by τ whose coefficients have size at most $2c(2N)^{10}$. The minimal equation for τ must divide this one, and so (with a new constant) $|D_\sigma| \le cN^{20}$ □

Theorem 9.9. *Assuming Conjectures* 7.8 *and* 9.5, *ZP holds for* $V_n \subset X_n$ *for all* n.

Remark 9.10. One might hope to prove at this juncture that Conjecture 9.5 implies finiteness of n-optimal tuples for each n (without assuming Conjecture 7.8). However, our proof will require Conjecture 7.8.

Proof. Suppose $C \in (\mathbb{C}^\times)^n$ is n-optimal. Then the point $(C, C) \in V_n$ is an optimal component. Thus C is grounded, and the relation group $\Gamma(C)$ has rank exceeding k, the number of cliques in C (because (C, C) must be an unlikely intersection).

By Lemma 9.7 we find a submodule Γ of relations on C with rank k and with $\Delta(\Gamma) \le c(n)\Gamma(C)$. By Conjecture 9.8 with $\epsilon = (20n)^{-1}$ we have

$$h(C) \le c(n)\Delta(C, \Gamma)^{1/(20n)}.$$

We now obtain a lower bound for $d(C)$ in terms of $\Delta(C)$. We start considering $\Delta_{\mathrm{mod}}(C)$. If x, y are distinct and $\Phi_N(x, y) = 0$ with N minimal then, as in the proof of Proposition 8.8,

$$N \le c(n)[\mathbb{Q}(x, y) : \mathbb{Q}]^4 \Delta(C, \Gamma)^{1/(10n)}.$$

If x is special, then we have $\Phi_N(x, x) = 0$ for some $N \le |D_x|$, while $d(x) \gg_\epsilon |D_x|^{1/2-\epsilon}$ (ineffectively) by Landau-Siegel. Thus again $N \le c[\mathbb{Q}(x) : \mathbb{Q}]^4$, and we find

$$\Delta_{\mathrm{mod}}(C) \le c(n)d(C)^4 \Delta(C, \Gamma)^{1/(10n)}$$

for some (ineffective if any c_i are special) positive $c(n)$.

Now by Proposition 9.6, there is a basis of $\Gamma(C)$ of height at most

$$\Delta(\Gamma(C)) \le c(n)d(C)^{4n}\Delta(C, \Gamma)^{1/10},$$

and since $\Delta(\Gamma) \le c(n)\Delta(C)$ we have that $\Delta(C) \le c(n)d(C)^{4n}\Delta(C)^{1/10}$. Hence

$$\Delta(C) \le c(n)d(C)^{5n}.$$

Now consider the uniformisation $\pi : F \to X$, which is definable in an o-minimal structure. As in the proof of 1.2, the point C and each of its conjugates gives rise to a rational point P of height $H(P) \leq c\Delta(C)$ on a suitable definable subset of a suitable power of $\mathrm{GL}_2(\mathbb{R})$. We follow the argument in the proof of Theorem 1.2 (which follows that in [13,14]). If Δ is sufficiently large then the Counting Theorem implies that the above-mentioned definable set contains positive-dimensional real semi-algebraic sets.

Since there are many conjugates of C giving rise to rational points, some positive-dimensional semi-algebraic set must give rise to a moving component of the given dimension and defect. Complexifying the real parameter, there is a larger component of V_n with the same defect. The mixed Ax-Schanuel implies there is a larger geodesic component with the same defect so that (in virtue of Conjecture 7.8) the point C was not n-optimal. This contradiction shows that the complexity of an optimal n-tuple is bounded. Then the degree $d(C)$ and the height $h(C)$ are bounded by some $c(n)$, and so there are only finitely many such C. $\square$

References

[1] J. AX, *On Schanuel's conjectures*, Ann. of Math. **93** (1971), 252–268.

[2] Y. BILU, F. LUCA and A. PIZARRO-MADARIAGA, *Rational products of singular moduli*, J. Number Theory **168** (2016), 397–410.

[3] Y. BILU, D. MASSER and U. ZANNIER, *An effective "theorem of André" for CM-points on a plane curve*, Math. Proc. Cambridge Philos. Soc. **154** (2013), 145–152.

[4] E. BOMBIERI, D. MASSER and U. ZANNIER, *Anomalous subvarieties – structure theorems and applications*, Int. Math. Res. Not. IMRN **19** (2007).

[5] C. DAW and A. HARRIS, *Categoricity of modular and Shimura curves*, J. Inst. Math. Jussieu **16** (2017), 1075–1101.

[6] N. ELKIES, *Shimura curve computations*, In: "Algorithmic Number Theory", Third International Symposium, ANTS-III, Portland, oregon, 1998, J. Buhler (ed.), Lecture Notes in Comp. Sci., Vol. 1423, Springer, Berlin, 1998, 1–47.

[7] E. ERRTHUM, *Singular moduli of Shimura curves*, Canad. J. Math. **63** (2011), 826–861.

[8] É. GAUDRON and G. RÉMOND, *Théorème des périodes et degrés minimaux d'isogénies*, Comment. Math. Helv. **89** (2014), 343–403.

[9] P. HABEGGER, *On the bounded height conjecture*, Int. Math. Res. Not. IMRN **2009**, 860–886.

[10] P. HABEGGER, *Weakly bounded height on modular curves*, Acta Math. Vietnam. **35** (2010), 43–69.

[11] P. HABEGGER, *Effective height upper bounds on algebraic tori*, CIRM Luminy course notes, arXiv:1201.3255.

[12] P. HABEGGER, *Singular moduli that are algebraic units*, Algebra Number Theory **9** (2015), 1515–1524.

[13] P. HABEGGER and J. PILA, *Some unlikely intersections beyond André–Oort*, Compos. Math. **148** (2012), 1–27.

[14] P. HABEGGER and J. PILA, *O-minimality and certain atypical intersections*, Ann. Sci. École Norm. Sup. (4) **49** (2016), 813–858.

[15] S. LANG, "Elliptic Functions", Graduate Texts in Mathematics **112**, second edition, Springer, New York, 1987.

[16] S. LOJASIEWICZ, "Introduction to Complex Analytic Geometry", Birkhäuser, Basel, 1991.

[17] T. LOHER and D. W. MASSER, *Uniformly counting points of bounded height*, Acta Arith. **111** (2004), 277–297.

[18] D. MASSER, *Linear relations on algebraic groups*, In: "New Advances in Transcendence Theory", Baker (ed.), CUP, 1988, 248–262.

[19] D. MASSER and G. WÜSTHOLZ, *Estimating isogenies on elliptic curves*, Invent. Math. **100** (1990), 1–24.

[20] R. PAULIN, *An explicit André-Oort type result for* $\mathbb{P}^1(\mathbb{C}) \times \mathbb{G}_m(\mathbb{C})$ *based on logarithmic forms*, Publ. Math. Debrecen **88** (2016), 21–33.

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

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

[23] J. PILA, *Functional transcendence via o-minimality, lecture notes, LMS-EPSRC mini-course*, 2013, In: "O-minimality and Diophantine Geometry", Jones and Wilkie (eds.), London Math. Soc. Lecture Note Series, Vol. 421, CUP, 2015.

[24] J. PILA and J. TSIMERMAN, *Ax-Schanuel for the* j*-function*, Duke Math. J. **165** (2016), 2587–2605.

[25] J. PILA and A. WILKIE, *The rational points of a definable set*, Duke Math. J. **133** (2006), 591–616.

[26] R. PINK, *A common generalization of the conjectures of André-Oort, Manin-Mumford, and Mordell-Lang*, 2005 preprint, available from the author's webpage.

[27] B. POIZAT, *L'egalité au cube*, J. Symbolic Logic **66** (2001), 1647–1676.

[28] M. RAYNAUD, *Hauteurs et isogénies*, Seminaire sur les pinceaux arithmetiques: La conjecture de Mordell, Astérisque **127** (1985), 199–234.

[29] J.-P. SERRE, "Lectures on the Mordell-Weil Theorem", *Aspects of Mathematics* E15, Vieweg, Braunschweig, 1989.

[30] G. SHIMURA, *Construction of class fields and zeta functions of algebraic curves*, Ann. of Math. **85** (1967), 58–159.

[31] J. SILVERMAN, *Heights and elliptic curves*, In: "Arithmetic Geometry", Cornell and Silverman (eds), Springer, New York, 1986, 253–265.

[32] J. TSIMERMAN, *Ax-Schanuel and o-minimality*, In: "O-minimality and Diophantine Geometry", Jones and Wilkie (eds.), London Math. Soc. Lecture Note Series, Vol. 421, CUP, 2015.

[33] D. ZAGIER, *Elliptic modular functions and their applications*, In: "The 1-2-3 of Modular Forms", J. Bruinier, G. van der Geer, G. Harder, and D. Zagier (eds.), Universitext, Springer, Berlin, 2008, 1–103.

[34] U. ZANNIER, "Some Problems of Unlikely Intersections in Arithmetic and Geometry", with appendices by D. Masser, Annals of Mathematics Studies, Vol. 181, Princeton University Press, 2012.

[35] B. ZILBER, *Exponential sums equations and the Schanuel conjecture*, J. Lond. Math. Soc. (2) **65** (2002), 27–44.

Mathematical Institute
University of Oxford
Woodstock Road
Oxford OX26GG, UK
pila@maths.ox.ac.uk

Department of Mathematics
University of Toronto
215 Huron Street
Toronto HU1001B, Canada
jacobt@math.toronto.edu