ACTA MATHEMATICA VIETNAMICA
Volume 35, Number 1, 2010, pp. 43–69

WEAKLY BOUNDED HEIGHT ON MODULAR CURVES

P. HABEGGER

Abstract. We study the intersection of a fixed plane algebraic curve C with modular curves of varying level. The height of points in such intersections cannot be bounded from above independently of the level when C is defined over the field of algebraic numbers. But we find a certain class of curves C for which the height is bounded logarithmically in the level. This bound is strong enough to imply certain finiteness result. Such evidence leads to a conjecture involving a logarithmic height bound unless C is of so-called special type. We also discuss connections to recent progress on conjectures concerning unlikely intersections.

1. Introduction

In this article we investigate points on a fixed plane algebraic curve whose coordinates are j-invariants of two varying isogenous elliptic curves. To this end we regard the affine line $\mathbf{A}^1$ over the field of complex numbers as the j-line. In other words it parameterizes elliptic curves defined over $\mathbf{C}$ up-to isomorphism. We identify all varieties with their set of complex points.

For an integer $N \geq 1$ let $\Phi_N \in \mathbf{Z}[X, Y]$ denote the Nth modular transformation polynomial and let $Y_0(N) \subset \mathbf{A}^2$ be the variety it cuts out. Then $Y_0(N)$ is a geometrically irreducible algebraic curve. Moreover, the coordinates of a point on this curve are j-invariants of two elliptic curves which are related by a cyclic isogeny of degree N. For more details we refer to Chapter 5 of [19].

Let $C \subset \mathbf{A}^2$ be an irreducible algebraic curve defined over $\mathbf{C}$.

We call a complex number a singular j-invariant if it is the j-invariant of an elliptic curve with complex multiplication. Singular j-invariants are algebraic integers. For the moment we assume that C contains infinitely many points (x, y) with x and y singular j-invariants. The André-Oort Conjecture predicts that C is either $Y_0(N)$ for some integer $N \geq 1$, or $\{j_0\} \times \mathbf{A}^1$, or $\mathbf{A}^1 \times \{j_0\}$ for some singular j-invariant j_0. This conjecture holds unconditionally for curves in $\mathbf{A}^2$ by a result of André [1].

The three types of curves

$$Y_0(N) \quad (N \geq 1), \quad \{j_0\} \times \mathbf{A}^1, \quad \mathbf{A}^1 \times \{j_0\} \quad (j_0 \text{ a singular } j\text{-invariant})$$

Received August 13, 2009; in revised form October 23, 2009.
2000 *Mathematics Subject Classification.* Primary: 11G50, Secondary: 11G18, 14G35.
Key words and phrases. Heights, modular curves, André-Oort Conjecture.

which arise above are called special curves. Together with

$$\mathbf{A}^2 \quad \text{and} \quad \{(j_0, j_0')\} \quad (j_0 \text{ and } j_0' \text{ singular } j\text{-invariants})$$

they are the class of special subvarieties of $\mathbf{A}^2$.

We will mainly be interested in the height of algebraic points in the intersection of C with $Y_0(N)$ as N varies. For such investigations it makes sense to suppose C is defined over $\overline{\mathbf{Q}}$, the algebraic closure of $\mathbf{Q}$ in $\mathbf{C}$. For a definition of the height we refer to Section 2.

After stating the main results we will motivate such questions by connecting them to general conjectures on unlikely intersections.

Our first result is a height upper bound which holds for a certain class of curves.

Theorem 1.1. *Let $C \subset \mathbf{A}^2$ be an irreducible algebraic curve defined over $\overline{\mathbf{Q}}$. Let X and Y denote the coordinate functions on $\mathbf{A}^2$ restricted to C. We assume $\deg X \neq \deg Y$. There exists a constant $c(C) \geq 0$ such that if $N \geq 1$ is an integer and $P \in (C \cap Y_0(N))(\overline{\mathbf{Q}})$, then*

$$h(P) \leq c(C) \log(1 + N). \tag{1.1}$$

The proof is quite short and based on the following idea: we interpret the coordinates of P as j-invariants of elliptic curves. A theorem of Faltings implies that both coordinates have nearly the same height. But coordinates of points on curves with $\deg X \neq \deg Y$ must have substantially different height by a result essentially going back to Néron and Siegel.

We believe that the condition $\deg X \neq \deg Y$ can be replaced by the assumption that C is not a special curve, cf. Theorem 1.3 and the conjecture below.

Let $C \subset \mathbf{A}^2$ be an irreducible curve that is not special and defined over $\overline{\mathbf{Q}}$. For any $\epsilon > 0$, one can apply Cohen's estimate [9] for the height of modular transformation polynomials to obtain a height bound of the form $c(C, \epsilon) N^{1+\epsilon}$ instead of the logarithmic estimate in (1.1). Here $c(C, \epsilon) \geq 0$ depends only on C and ϵ but not on N.

Can the upper bound in (1.1) be replaced by non-decreasing function that is $o(\log N)$ for large N? The answer is no by the following theorem.

Theorem 1.2. *Let $C \subset \mathbf{A}^2$ be an irreducible algebraic curve defined over $\overline{\mathbf{Q}}$ that is not a special curve. There are constants $c(C) > 0$ and $p_0(C)$ such that if p is a prime with $p \geq p_0(C)$ then there exists a point $P \in (C \cap Y_0(p))(\overline{\mathbf{Q}})$ with*

$$h(P) \geq c(C) \log p.$$

The proof of Theorem 1.2 is more involved than the proof of Theorem 1.1. We need a recent lower bound for linear forms in elliptic logarithms by David and Hirata-Kohno [13]. A further ingredient is an equidistribution result for Hecke orbits by Clozel and Ullmo [7]. Finally, the method involves an auxiliary point whose existence is guaranteed by the André-Oort Conjecture. As mentioned above, this conjecture is known in our situation.

Let

$$\mathcal{S}' = \bigcup_{N \geq 1} Y_0(N). \tag{1.2}$$

An immediate conclusion of the last theorem is that $C(\overline{\mathbf{Q}}) \cap \mathcal{S}'$ is infinite and has unbounded height.

By Northcott's Theorem, a subset of $C(\overline{\mathbf{Q}})$ of bounded height and bounded degree is finite. Because the height is not uniformly bounded on $C(\overline{\mathbf{Q}}) \cap \mathcal{S}'$, Northcott's Theorem cannot be used to show that a subset of bounded degree is finite. Nevertheless, in the corollary below we do obtain finiteness results in this spirit. Instead of using Northcott's Theorem we apply Masser and Wüstholz's isogeny estimates for elliptic curves [21].

Corollary 1.1. *Let C be as in Theorem 1.1. For any real D the set*

$$\{P \in C(\overline{\mathbf{Q}}) \cap \mathcal{S}';\ [\mathbf{Q}(P) : \mathbf{Q}] \leq D\}$$

is finite.

We give an example of a curve which does not satisfy the hypothesis of Theorem 1.1 but for which the conclusion, at least for prime level, holds. The proof involves, among other things, Cohen's estimate mentioned above.

Theorem 1.3. *Let $C \subset \mathbf{A}^2$ be the curve defined by $X - Y = 1$. There exists an absolute constant c such that if p is a prime and $P \in (C \cap Y_0(p))(\overline{\mathbf{Q}})$, then*

$$h(P) \leq 6 \log(p) + c.$$

Unfortunately, the proof cannot treat other curves such as $X + Y = 1$; cf. the end of section 6.

A finiteness result analogous to Corollary 1.1 for the curve given by $X - Y = 1$ can also be proven on restricting the union (1.2) to prime level. But the argument given in the proof of Theorem 1.3 shows $[\mathbf{Q}(P) : \mathbf{Q}] \geq p$ directly without appealing to isogeny estimates.

We now put the above results into broader context. Pink [25] stated a vast generalization of the André-Oort Conjecture. His Conjecture 1.3 governs the intersection of a subvariety Z of a mixed Shimura variety with the union of all special subvarieties of codimension strictly greater than $\dim Z$. Accordingly, this intersection should not be Zariski dense unless Z is contained in a special subvariety of positive codimension inside the ambient Shimura variety.

Our terminology of special subvarieties is consistent with the language of Shimura varieties when considering $\mathbf{A}^2$ as a Shimura variety, cf. Proposition 2.1 [15]. Pink's Conjecture makes no prediction on the intersection of C with varying $Y_0(N)$ because the codimension of $Y_0(N)$ equals the dimension of C. On the other hand, recent progress made on unlikely intersections of a fixed subvariety of a semi-abelian variety with varying algebraic subgroups suggests studying heights on intersections as in our situation.

More precisely, let us for the moment assume that C is an irreducible algebraic curve contained in the algebraic torus $\mathbf{G}_m^n$. We assume that this curve is not

contained in the translate of a proper algebraic subgroup of $\mathbf{G}_m^n$. Bombieri, Masser, and Zannier [3] proved the following result: there exists B such that any point of C contained in the union over all algebraic subgroups of codimension at least 1 has height at most B. This is of course in contrast to the modular situation where the height is unbounded. They then applied this height bound and other tools to prove the following statement: only finitely many points on C are contained in the union of all algebraic subgroups of codimension at least 2.

Following this general strategy of using a height bound to prove finiteness, Maurin [22] was able to weaken the hypothesis on C. His theorem is the precise analog of Pink's Conjecture for curves defined over $\overline{\mathbf{Q}}$ in the algebraic torus.

We return to the modular setting. In view of Pink's Conjecture it seems natural to study the intersection of our curve $C \subset \mathbf{A}^2$ with the union of all special curves in $\mathbf{A}^2$ and not just the modular curves $Y_0(N)$. Bombieri, Masser, and Zannier [4] noticed that singular j-invariants have unbounded height. Even earlier, Colmez [10] proved this by supplying a height lower bound. Both statements imply that the set of points on C having a coordinate that is a singular j-invariant has unbounded height.

Let us define

$$\mathcal{S} = \mathcal{S}' \cup \bigcup_{\substack{j_0 \text{ singular} \\ j\text{-invariant}}} (\{j_0\} \times \mathbf{A}^1) \cup (\mathbf{A}^1 \times \{j_0\}).$$

This set is the union of all special subvarieties of $\mathbf{A}^2$ of codimension at least 1; it contains all points whose coordinates are both singular j-invariants.

There are only finitely many singular j-invariants of bounded degree. So Corollary 1.1 remains true with $\mathcal{S}'$ replaced by $\mathcal{S}$, at least if we assume that C is not special.

In view of Theorem 1.1 we can hope for a good height bound for $P \in C(\overline{\mathbf{Q}}) \cap \mathcal{S}$ in terms of the minimal special variety containing P. We will soon formulate a Conjecture on Weakly Bounded Height for such points. But first we need to introduce some notation.

For $P \in \mathbf{A}^2$ we define the Zariski closed set

$$\mathcal{S}(P) = \bigcap_{\substack{S \subset \mathbf{A}^2 \text{ special}, \\ P \in S}} S.$$

We claim that $\mathcal{S}(P)$ is itself special. Indeed, say $P = (x, y)$ and let us assume $\mathcal{S}(P) \neq \mathbf{A}^2$. If x and y are both singular j-invariants, then $\mathcal{S}(P) = \{P\}$. If precisely one among x and y is a singular j-invariant, say x, then $\mathcal{S}(P) \subset \{x\} \times \mathbf{A}^1$. We cannot have $\mathcal{S}(P) \subset Y_0(N)$ for some $N \geq 1$, since otherwise y would be a singular j-invariant too. Hence $\mathcal{S}(P)$ cannot lie in any special subvariety of codimension 1 other than $\{x\} \times \mathbf{A}^1$. We get $\mathcal{S}(P) = \{x\} \times \mathbf{A}^1$. Finally, if neither x or y are singular j-invariants, then P is on $Y_0(N)$ for some $N \geq 1$. Lemma 3.2 below implies that if $P \in Y_0(N) \cap Y_0(M)$ for an integer M, then $N = M$. We conclude $\mathcal{S}(P) = Y_0(N)$.

Our claim follows and we have showed that if $\mathcal{S}(P) \neq \mathbf{A}^2$, then

(1.3)
$$\mathcal{S}(P) = \begin{cases} \{P\} & : \text{ if } x \text{ and } y \text{ are singular } j\text{-invariants,} \\ \{x\} \times \mathbf{A}^1 & : \text{ if } x \text{ is a singular } j\text{-invariant and } y \text{ is not,} \\ \mathbf{A}^1 \times \{y\} & : \text{ if } y \text{ is a singular } j\text{-invariant and } x \text{ is not,} \\ Y_0(N) & : \text{ if } P \in Y_0(N) \text{ and } x,y \text{ are not singular } j\text{-invariants.} \end{cases}$$

Of course, $\mathcal{S}(P)$ is always defined over $\overline{\mathbf{Q}}$. But it needs not be defined over the rationals. We let $\mathcal{S}_{\mathbf{Q}}(P) \subset \mathbf{A}^2$ denote the union of all conjugates over $\mathbf{Q}$ of $\mathcal{S}(P)$. It is the minimal variety defined over $\mathbf{Q}$ containing $\mathcal{S}(P)$.

We fix the open immersion $\mathbf{A}^2 \to \mathbf{P}^1 \times \mathbf{P}^1$ that maps (x, y) to $([x : 1], [y : 1])$. Let us consider the pull-backs of the ample generator of the Picard group of $\mathbf{P}^1$ by the two projections $\mathbf{P}^1 \times \mathbf{P}^1 \to \mathbf{P}^1$. Their tensor product is an ample line bundle on $\mathbf{P}^1 \times \mathbf{P}^1$. We use this line bundle to define the degree of a Zariski closed set in $\mathbf{P}^1 \times \mathbf{P}^1$. Let $Z \subset \mathbf{A}^2$ be Zariski closed, we let $\deg Z$ denote the degree of the Zariski closure of Z in $\mathbf{P}^1 \times \mathbf{P}^1$.

For example, $\deg \{x\} \times \mathbf{A}^1 = 1$. It follows from Theorem 3, page 55 of [19] that $\deg_X \Phi_N = \deg_Y \Phi_N = \Psi(N)$ where $\Psi(N) = N \prod_l (1 + l^{-1})$ and the product runs over all primes l dividing N. So $\deg Y_0(N) = 2\Psi(N)$. For $P \in \mathcal{S}$ we have

(1.4)
$$\deg \mathcal{S}_{\mathbf{Q}}(P) = \begin{cases} [\mathbf{Q}(x, y) : \mathbf{Q}] & : \text{ if } x \text{ and } y \text{ are singular } j\text{-invariants,} \\ [\mathbf{Q}(x) : \mathbf{Q}] & : \text{ if } x \text{ is a singular } j\text{-invariant and } y \text{ is not,} \\ [\mathbf{Q}(y) : \mathbf{Q}] & : \text{ if } y \text{ is a singular } j\text{-invariant and } x \text{ is not,} \\ 2\Psi(N) & : \text{ if } P \in Y_0(N) \text{ and } x,y \text{ are not singular } j\text{-invariants.} \end{cases}$$

We now state our Conjecture on Weakly Bounded Height.

Conjecture. *Let $C \subset \mathbf{A}^2$ be an irreducible algebraic curve defined over $\overline{\mathbf{Q}}$ that is not special. There exists a constant $c(C) \geq 0$ such that if $P \in C(\overline{\mathbf{Q}}) \cap \mathcal{S}$, then*

$$h(P) \leq c(C) \log(1 + \deg \mathcal{S}_{\mathbf{Q}}(P)).$$

Unfortunately, we are not able to prove this conjecture for the class of curves appearing in Theorem 1.1. Below we will show it for these curves in a weaker form. In fact it is the special curves $\{j_0\} \times \mathbf{A}^1$ and $\mathbf{A}^1 \times \{j_0\}$ we forgot in Theorem 1.1 which cause additional trouble. These problems disappear if we assume the Generalized Riemann Hypothesis (GRH).

Corollary 1.2. *Let $C \subset \mathbf{A}^2$ be as in Theorem 1.1 and let us also assume that C is not special.*

(i) *There exists a constant $c(C) \geq 0$ such that if $P \in C(\overline{\mathbf{Q}}) \cap \mathcal{S}'$, then*

$$h(P) \leq c(C) \log(1 + \deg \mathcal{S}_{\mathbf{Q}}(P)).$$

(ii) *If the GRH holds for all odd, real Dirichlet L-functions, then the Conjecture on Weakly Bounded Height holds for C.*

2. Preliminary remarks and notations

We begin by defining a height function needed throughout this article.

Let K be a number field, in other words a finite extension of $\mathbf{Q}$. We let M_K^0 denote the set of non-trivial, non-archimedean absolute values on K whose restriction to $\mathbf{Q}$ is the p-adic absolute value for some prime p. For $v \in M_K^0$ we set d_v to be the degree of the completion of K with respect to v over the completion of $\mathbf{Q}$ with respect to the restriction of v to $\mathbf{Q}$.

We define the height of $x \in K$ to be

(2.1)
$$h(x) = \frac{1}{[K:\mathbf{Q}]} \left(\sum_{\sigma: K \to \mathbf{C}} \log \max\{1, |\sigma(x)|\} + \sum_{v \in M_K^0} d_v \log \max\{1, |x|_v\} \right) \geq 0,$$

here and throughout this article $\sum_{\sigma:K\to\mathbf{C}}$ means taking the sum over the $[K:\mathbf{Q}]$ different embeddings $K \to \mathbf{C}$. Our height does not change when replacing K by another number field containing x. Hence it determines a well-defined function $h : \overline{\mathbf{Q}} \to [0, \infty)$. One may show that it is invariant under the action on $\overline{\mathbf{Q}}$ of the absolute Galois group of $\mathbf{Q}$. A reference for heights is Chapter 1 of [2].

From the point of view of notation, it is sometimes useful to work with the exponential height $H(x) = \exp h(x) \geq 1$.

For $(x, y) \in \mathbf{A}^2(\overline{\mathbf{Q}})$ we define $h(x, y) = \max\{h(x), h(y)\}$.

Some properties our height satisfies are $H(xy) \leq H(x)H(y)$, $H(x + y) \leq 2H(x)H(y)$ for $x, y \in \overline{\mathbf{Q}}$, $H(x) = \max\{1, |x|\}$ for integers x, and $H(x^n) = H(x)^{|n|}$ for $x \in \overline{\mathbf{Q}}\backslash\{0\}$ and $n \in \mathbf{Z}$. We shall refer to these as elementary height properties. The first three may be deduced directly from (2.1) and for the last one needs the product formula $\sum_{\sigma:K\to\mathbf{C}} \log |\sigma(x)| + \sum_{v\in M_K^0} d_v \log |x|_v = 0$ which holds for all non-zero $x \in K$.

If K is an arbitrary number field, Δ_K will denote its discriminant, h_K its class number, and $\mathcal{O}_K$ its ring of algebraic integers. If K is quadratic we let $\chi_K(\cdot) = \left(\frac{\Delta_K}{\cdot}\right)$ denote the associated Dirichlet character where $\left(\frac{\cdot}{\cdot}\right)$ is the Kronecker symbol.

For any elliptic curve E defined over $\overline{\mathbf{Q}}$ we let $h_F(E)$ denote its absolute stable Faltings height [16, 27]. We will use Deligne's normalization [14]. Hence our Faltings height equals the one given in Silverman's article [27] minus a universal constant.

3. Proof of Theorem 1.1 and Corollary 1.1

We begin by a simple lemma on elliptic curves.

Lemma 3.1. *Let E be an elliptic curve defined over $\overline{\mathbf{Q}}$, let j be its j-invariant, and let K be a number field containing j. There exist $g_2, g_3 \in K$ such that E is isomorphic to the elliptic curve defined by the Weierstrass equation $Y^2 = 4X^3 - g_2X - g_3$ and such that*

$$\max\{h(g_2), h(g_3)\} \leq c_1(1 + h(j));$$

here c_1 is an absolute constant.

Proof. We note $Y^2 = 4X^3 - 4$ defines an elliptic curve with j-invariant 0 and $Y^2 = 4X^3 - 4X$ defines an elliptic curve with j-invariant 1728. In both cases $\max\{h(g_2), h(g_3)\} = \log 4 \leq (\log 4)(1 + h(j))$. Hence it suffices to prove the lemma if $j \neq 0, 1728$.

In this case we choose $g_2 = j/(12(j - 1728))$ and $g_3 = -j/(6^3(j - 1728))$. A direct calculation shows that $Y^2 = 4X^3 - g_2X - g_3$ defines an elliptic curve with j-invariant equal to j. By the elementary height properties we see $h(g_2) \leq h(j) + h(12(j - 1728)) \leq h(j) + h(12) + h(j - 1728) \leq \log(2^9 3^4) + 2h(j)$ and $h(g_3) \leq h(j) + h(6^3(j - 1728)) \leq h(j) + h(6^3) + h(j - 1728) \leq \log(2^{10}3^6) + 2h(j)$. □

We will also need the following simple statement for elliptic curves without complex multiplication. An isogeny between two elliptic curves is called cyclic if it has cyclic kernel.

Lemma 3.2. *Let E and E' be elliptic curves defined over $\mathbf{C}$ and assume E has endomorphism ring $\mathrm{End}(E) = \mathbf{Z}$. Let $\varphi : E \to E'$ be a cyclic isogeny. Then φ generates the group of homomorphisms $\mathrm{Hom}(E, E')$.*

Proof. Let $\hat{\varphi} : E' \to E$ be the dual of φ. For $\psi \in \mathrm{Hom}(E, E')$ we have $\hat{\varphi} \circ \psi \in \mathrm{End}(E)$. This construction gives an injective homomorphism $\mathrm{Hom}(E, E') \to \mathrm{End}(E)$. It follows from our hypothesis that $\mathrm{Hom}(E, E')$ has rank at most 1. But it must be infinite cyclic because it contains φ. If ψ is a generator of $\mathrm{Hom}(E, E')$, then $\varphi = \lambda\psi$ for some $\lambda \in \mathbf{Z}$. We conclude that the kernel of multiplication by λ is isomorphic to a quotient of the kernel of φ; thus itself cyclic. In characteristic 0 this is only possible if $\lambda = \pm 1$. □

To prove Theorem 1.1 we need to compare the Faltings height with the height of the j-invariant. Silverman's Proposition 2.1 implies

$$h(j_E) - c_2 \log(3 + h(j_E)) \leq 12h_F(E) \leq h(j_E) + c_2 \tag{3.1}$$

if j_E is the j-invariant of an elliptic curve E defined over $\overline{\mathbf{Q}}$; here $c_2 > 0$ is an absolute constant. In Silverman's work the unstable minimal discriminant is the unit ideal because when defining the stable Faltings height we assume that E has semi-stable reduction everywhere.

We now prove Theorem 1.1.

Say d_1 and d_2 are the degrees of the first, respectively the second coordinate function on C. A theorem of Néron [24] implies that there is a constant $c_3 > 0$ depending on C but not on x or y such that

$$|h(x)d_2 - h(y)d_1| \leq c_3(1 + h(x))^{1/2} \tag{3.2}$$

for all $(x, y) \in C(\overline{\mathbf{Q}})$. The degrees d_1 and d_2 are not both zero, without loss of generality we may assume $d_1 \geq 1$. In particular,

$$h(y) \leq h(x)d_2 + c_3(1 + h(x))^{1/2}. \tag{3.3}$$

We suppose $(x,y) \in (C \cap Y_0(N))(\overline{\mathbf{Q}})$ for some $N \geq 1$. Let E_1 and E_2 be elliptic curves with j-invariants x and y, respectively. There is a cyclic isogeny of degree N between E_1 and E_2.

The behavior of the Faltings height under isogenies is well-studied. For example by the original work of Faltings, Lemma 5 of [16], or by Corollary 2.1.4 of [26] we have

$$|h_F(E_1) - h_F(E_2)| \leq \frac{1}{2}\log N.$$

Combining this bound with (3.1) we get

$$\begin{aligned} |h(y) - h(x)| &\leq |h(x) - 12h_F(E_1)| + |h(y) - 12h_F(E_2)| + 12|h_F(E_1) - h_F(E_2)| \\ &\leq 2c_2 \log(3 + h(x) + h(y)) + 6\log N \\ &\leq c_3 \log(3 + h(x)) + 6\log N \end{aligned} \tag{3.4}$$

in the last inequality we used the fact that one can bound $h(y)$ from above in terms of $h(x)$, this follows from (3.3).

Since $d_1 \neq d_2$ we have $h(x) \leq h(x)|d_2 - d_1| \leq |h(x)d_2 - h(y)d_1| + |h(y) - h(x)|d_1$, hence

$$h(x) \leq c_4(1 + h(x))^{1/2} + 6d_1 \log N$$

from (3.2) and (3.4) with $c_4 > 0$ independent of N and x. We conclude $h(x) \leq c_5 \log(1+N)$ with $c_5 > 0$ independent of N. We can find a similar bound for $h(y)$ using (3.3). □

We can now prove Corollary 1.1:

Since there are only finitely many singular j-invariants of bounded degree over $\mathbf{Q}$ it suffices to show that

$$\begin{aligned} \{(x,y) \in C(\overline{\mathbf{Q}}) \cap \mathcal{S}';\ &\text{neither } x \text{ or } y \text{ is a singular } j\text{-invariant} \\ &\text{and } [\mathbf{Q}(x,y) : \mathbf{Q}] \leq D\} \end{aligned} \tag{3.5}$$

is finite.

Let (x,y) be in this set, so it lies in $Y_0(N)(\overline{\mathbf{Q}})$ for some $N \geq 1$. Let E_1 and E_2 be the corresponding elliptic curves; they do not have complex multiplication. Moreover, there is a cyclic isogeny $\varphi : E_1 \to E_2$ of degree N.

We now consider an isogeny $\varphi' : E_1 \to E_2$ of minimal degree N'. By the isogeny estimates of Masser-Wüstholz [21], we have $N' \leq c'(D)(1+h(x))^4$ where $c'(D)$ is a constant depending only on D. We remark that Lemma 3.1 applied to E_1 enables us to compare Masser and Wüstholz's height of an elliptic curve with $h(x)$.

By Lemma 3.2 we know that φ generates $\mathrm{Hom}(E_1, E_2)$, so $\varphi' = \lambda\varphi$. Since φ' has minimal degree we must have $\lambda = \pm 1$. Hence $N' = N$ and we obtain

$$N \leq c'(D)(1 + h(x))^4. \tag{3.6}$$

Now Theorem 1.1 tells us that $h(x) \leq c\log(1+N)$ where c depends only on the curve C but not on N. On inserting this bound into (3.6) we conclude that N is bounded in terms of the curve C and the degree bound D.

Therefore, any point in (3.5) is contained in the intersection of C with a finite union of $Y_0(N)$.

We note that $\deg X \neq \deg Y$ implies $C \neq Y_0(N)$ for all $N \geq 1$. Indeed, the polynomial Φ_N defining $Y_0(N)$ has equal degree in both variables. So, $C \cap Y_0(N)$ is finite for all $N \geq 1$; the corollary follows. □

4. Height of elliptic curves with complex multiplication

Before we come to the proof of Theorem 1.2 we give some estimates on the Faltings height of an elliptic curve with complex multiplication. These will be used in the proof of Theorem 1.2 and Corollary 1.2. Some of these estimates are implicitly contained in the work of Colmez [10].

Let E be an elliptic curve with complex multiplication. So $\mathrm{End}(E)$ is an order in an imaginary quadratic number field L. For brevity we write $\mathcal{O} = \mathcal{O}_L, \Delta = \Delta_L$, and $\chi = \chi_L$. We note $|\Delta| \geq 3$. There exists a unique integer $f \geq 1$ such that $\mathrm{End}(E) = \mathbf{Z} + f\mathcal{O}$. For a prime p dividing f, or $p|f$ for brevity, we set

$$e_f(p) = \frac{1-\chi(p)}{p-\chi(p)} \frac{1-p^{-n}}{1-p^{-1}}$$

with n the maximal integer such that $p^n|f$.

We attach the Dirichlet L-function $L(\chi, s)$ to the odd, real character χ. It is entire and does not vanish at $s = 1$.

The first lemma expresses the Faltings height of E in terms of Δ and f.

Lemma 4.1. *We have*

$$2h_F(E) = \frac{1}{2}\log(|\Delta|f^2) + \frac{L'(\chi,1)}{L(\chi,1)} - \left(\sum_{p|f} e_f(p)\log p\right) - \gamma - \log(2\pi) \tag{4.1}$$

where γ is Euler's constant.

Proof. The main result of Nakkajima and Taguchi's work on the Chowla-Selberg formula [23] together with their Lemma 3 implies

$$\exp(2h_F(E)) = |\Delta|^{1/2} f \left(\prod_{a=1}^{|\Delta|} \Gamma\left(\frac{a}{|\Delta|}\right)^{\chi(a)}\right)^{-\frac{\#\mathcal{O}^\times}{2h_L}} \prod_{p|f} p^{-e_f(p)}, \tag{4.2}$$

here Γ denotes the gamma function.

We use Proposition 10.3.5 of [8] and the fact that $L(\chi,0) = |\Delta|^{1/2}L(\chi,1)/\pi = 2h_L/\#\mathcal{O}^\times$ (cf. Theorem 10.2.14 and Proposition 10.5.10 of [8]) to obtain

$$\sum_{a=1}^{|\Delta|} \chi(a)\log\Gamma\left(\frac{a}{|\Delta|}\right) = \frac{2h_L}{\#\mathcal{O}^\times}\left(-\frac{L'(\chi,1)}{L(\chi,1)} + \gamma + \log(2\pi)\right).$$

The present lemma follows from inserting this into the logarithm of the right-hand side of (4.2). □

If we fix the discriminant Δ and let f vary we circumvent the problem of controlling the logarithmic derivative of $L(\chi, s)$ in (4.1). The following estimate will be useful in the proof of Theorem 1.2.

Lemma 4.2. *Let Δ be fixed, for $f \geq 1$ we have*

$$h_F(E) = \frac{1}{2}\log f + O(\log\log(2+f))$$

where the implied constant is independent of f but may depend on Δ.

Proof. By virtue of Lemma 4.1 and since $\frac{L'(\chi,1)}{L(\chi,1)}$ is independent of f it suffices to show

$$\sum_{p|f} e_f(p)\log p = O(\log\log f) \tag{4.3}$$

for large f.

The left-hand side is certainly never negative. Furthermore, using the definition we estimate $e_f(p) \leq \frac{2}{p+1}\frac{1}{1-p^{-1}} = \frac{2}{p-p^{-1}}$. A short calculation shows $e_f(p) \leq 3/p$. We obtain

$$\sum_{p|f} e_f(p)\log p \leq 3\sum_{\substack{p|f\\ p\leq \log f}} \frac{\log p}{p} + 3\sum_{\substack{p|f\\ p>\log f}} \frac{\log p}{p}.$$

The first sum on the right is at most $O(\log\log f)$ by an equality on page 57, Chapter 7 of [12]. It remains to find such a bound for the second sum. As the function $p \mapsto \log(p)/p$ decreases for large p, we have

$$\sum_{\substack{p|f\\ p>\log f}} \frac{\log p}{p} \leq \frac{\log\log f}{\log f}\sum_{\substack{p|f\\ p>\log f}} 1.$$

The lemma follows since f has in total at most $\frac{\log f}{\log 2}$ distinct prime factors. □

Obtaining estimates that are uniform in $|\Delta|$ involves studying the logarithmic derivative of $L(\chi, s)$ at $s = 1$. We have a double logarithmic bound when assuming the GRH.

Lemma 4.3. *Assume the GRH holds for all real, odd Dirichlet L-functions. Then*

$$\frac{L'(\chi,1)}{L(\chi,1)} = O(\log\log|\Delta|)$$

where the implied constant is absolute.

Proof. The proof is standard in analytic number theory. For a proof one can consult Section 3.1 of [17]. □

Under the GRH and using the results from this section, one immediately verifies

$$h_F(E) = \frac{1}{4}\log|\Delta| + O(\log\log|\Delta|) \quad \text{for } f = 1.$$

For general $f \geq 1$, estimate (4.3) can be used to obtain

$$h_F(E) = \frac{1}{4}\log(|\Delta|f^2) + O(\log\log(|\Delta|f^2)).$$

The discriminant of the order $\mathrm{End}(E)$, denoted here by $\mathrm{Disc}(\mathrm{End}(E))$, has absolute value $|\Delta|f^2 \geq 3$. We obtain the asymptotic formula

$$h_F(E) = \frac{1}{4}\log|\mathrm{Disc}(\mathrm{End}(E))| + O(\log\log|\mathrm{Disc}(\mathrm{End}(E))|).$$

5. Unboundedness of height

In this section we prove Theorem 1.2.

Let $\mathbf{H}$ denote the complex upper half plane and let $j : \mathbf{H} \to \mathbf{C}$ be the elliptic modular function. We need to work with an auxiliary point $(j_1, j_2) \in \mathbf{A}^2(\overline{\mathbf{Q}})$ satisfying the following properties: j_1 is a singular j-invariant. Hence $j_1 = j(\tau)$ for an imaginary quadratic τ in $\mathcal{F}$, the usual (closed) fundamental domain for the action of $\mathrm{SL}_2(\mathbf{Z})$ on $\mathbf{H}$ given by fractional linear transformations. Furthermore, we assume that j_2 is algebraic but not a singular j-invariant. Since j_1 is algebraic we may fix a number field K containing both it and j_2.

Throughout this section p denotes a prime. We will work with positive constants $c_1, c_2, \ldots$ and $\delta \in (0, 1]$ that, if not stated otherwise, may depend on j_1, j_2, and K but are independent of p. Moreover, the c_i will not depend on δ either; in the end we will choose δ to depend on some c_i. Finally, we often need to work with an embedding of a fixed, finite extension of K into $\mathbf{C}$; we silently assume that the c_i and δ are independent of such an embedding.

We call p large enough if it is larger than some constant which may depend on δ, j_1, j_2, and K.

The norm $|\beta|$ of $\beta \in \mathrm{GL}_2(\mathbf{Q})$ is the largest absolute value of any of its entry. If $\rho \in \mathbf{H}$, then $\beta\rho$ denotes the usual action of β on ρ.

Let us define $p + 1$ square matrices

$$\alpha_0 = \begin{bmatrix} p & 0 \\ 0 & 1 \end{bmatrix}, \alpha_1 = \begin{bmatrix} 1 & 0 \\ 0 & p \end{bmatrix}, \alpha_2 = \begin{bmatrix} 1 & 1 \\ 0 & p \end{bmatrix}, \ldots, \alpha_p = \begin{bmatrix} 1 & p-1 \\ 0 & p \end{bmatrix}.$$

5.1. **A lower bound.** We need a preliminary lemma on moving imaginary quadratic elements of $\mathbf{H}$ into $\mathcal{F}$.

Lemma 5.1. *There are absolute constants c_1 and c_2 with the following properties: if $\rho \in \mathbf{H}$ is imaginary quadratic then $h(\beta\rho) \leq c_1(1 + \log|\beta| + h(\rho))$ for any $\beta \in \mathrm{SL}_2(\mathbf{Z})$. Moreover, there exists $\beta \in \mathrm{SL}_2(\mathbf{Z})$ with $\beta\rho \in \mathcal{F}$ and $h(\beta\rho) \leq c_2(1+h(\rho))$.*

Proof. The first statement is an easy consequence of elementary height properties. Indeed, say

$$\beta = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbf{Z}). \tag{5.1}$$

Then

$$\begin{aligned} H(\beta\rho) &= H((a\rho + b)/(c\rho + d)) \\ &\leq H(a\rho + b)H((c\rho + d)^{-1}) \\ &= H(a\rho + b)H(c\rho + d), \end{aligned}$$

giving

$$H(\beta\rho) \leq 4H(a)H(b)H(c)H(d)H(\rho)^2 \leq 4\max\{|a|,|b|,|c|,|d|\}^4 H(\rho)^2.$$

So $H(\beta\rho) \leq 4|\beta|^4 H(\rho)^2$ and thus $c_1 = 4$ becomes a valid choice.

For the second statement we write $\rho = x+yi$ with x, y real algebraic and $y > 0$. We note that $x = (\rho+\overline{\rho})/2$ and $y = (\rho-\overline{\rho})/2$ where $\overline{\rho}$ is the complex conjugate of ρ. Hence $H(x) \leq H(\rho+\overline{\rho})H(2) \leq 4H(\rho)H(\overline{\rho}) = 4H(\rho)^2$ and similarly $H(y) \leq 4H(\rho)^2$. We also remark that $|x| \leq H(x)$ and $y^{-1} \leq H(y^{-1})^2 = H(y)^2$ by elementary height properties and because x is rational and y has degree at most 2.

Let $im(\cdot)$ be the imaginary part of a complex number. In the first step we will show that there exists $\beta \in \mathrm{SL}_2(\mathbf{Z})$ with $im(\beta\rho) > 1/2$ such that $H(\beta\rho)$ is suitably bounded, i.e. it is bounded from above by a fixed power of $2H(\rho)$. We may certainly assume that $y \leq 1/2$. We set $Q = y^{-1/2} > 1$. By a theorem of Dirichlet there are integers c, d with $1 \leq c < Q$ such that $|cx + d| \leq Q^{-1}$, cf. page 1 of [6]. Without loss of generality we may assume that c and d are coprime. By the previous paragraph we have $|c| < Q = y^{-1/2} \leq H(y) \leq 4H(\rho)^2$ and $|d| \leq Q^{-1} + |cx| \leq 1 + QH(x) \leq 1 + 16H(\rho)^4 \leq 17H(\rho)^4$. Moreover, one may find integers a, b such that $ad - bc = 1$ with $\max\{|a|,|b|\}$ bounded by an absolute constant times $\max\{|c|,|d|\}$. The matrix β defined as in (5.1) is in $\mathrm{SL}_2(\mathbf{Z})$ and has norm bounded by a fixed power of $2H(\rho)$. So $H(\beta\rho)$ is suitably bounded by the first part of the lemma. Finally, $im(\beta\rho) = y/|c\rho + d|^2$ and $|c\rho + d|^2 = (cx + d)^2 + (cy)^2 < Q^{-2} + Q^2y^2 = 2y$. We arrive at $im(\beta\rho) > 1/2$.

In the second step we will show that for any imaginary quadratic $\rho \in \mathbf{H}$ there is $\beta \in \mathrm{SL}_2(\mathbf{Z})$ such that $\beta\rho$ has real part in $[-1/2, 1/2]$ and such that $H(\beta\rho)$ is suitably bounded. It suffices to choose β with $a = d = 1$, $c = 0$, and b an integer with $|x + b| \leq 1/2$. We note $|\beta| \leq \max\{1, |x| + 1/2\} \leq 1 + 4H(\rho)^2 \leq 5H(\rho)^2$. So $H(\beta\rho)$ is suitably bounded by the first part of the lemma.

The third and last step consists in showing that if $|x| \leq 1/2$ and $y > 1/2$ there is $\beta \in \mathrm{SL}_2(\mathbf{Z})$ such that $\beta\rho \in \mathcal{F}$ and such that $H(\beta\rho)$ is suitably bounded. It is not difficult to verify that the heights of ρ, $-1/\rho$, or $\pm 1 - 1/\rho$ are suitably bounded and that at least one lies in $\mathcal{F}$.

The proof of the second part of the lemma follows from these three steps. □

We recall $j_1 = j(\tau)$. We may factor

$$\Phi_p(j_1, Y) = \prod_{k=0}^{p} (Y - j(\alpha_k\tau)) \in K[Y], \tag{5.2}$$

for a reference see §2, Chapter 5 of [19]. If $\sigma : K \to \mathbf{C}$ is an embedding we choose $\tau_\sigma \in \mathcal{F}$ with $j(\tau_\sigma) = \sigma(j(\tau)) = \sigma(j_1)$. Let us extend σ to the number field $K(j(\alpha_0\tau), \ldots, j(\alpha_p\tau))$ and apply it in the usual manner to polynomials with coefficients in this field. Since Φ_p has integer coefficients we obtain

$$\Phi_p(\sigma(j_1), Y) = \sigma(\Phi_p(j_1, Y)) = \prod_{k=0}^{p} (Y - \sigma(j(\alpha_k\tau))). \tag{5.3}$$

One the other hand, using $\sigma(j_1) = j(\tau_\sigma)$ and (5.2) with τ_σ instead of τ yields

$$\Phi_p(\sigma(j_1), Y) = \prod_{k=0}^{p} (Y - j(\alpha_k \tau_\sigma)).$$

Therefore, there exists a permutation of $\{0, \dots, p\}$, which we also denote with σ, such that

$$\sigma(j(\alpha_k \tau)) = j(\alpha_{\sigma(k)} \tau_\sigma) \quad \text{for} \quad 0 \le k \le p.$$

We come to a first lower bound. It uses a recent lower bound for linear forms in elliptic logarithms by David and Hirata-Kohno [13].

Lemma 5.2. *We have $j(\alpha_k \tau) \neq j_2$ and*

$$\log |\sigma(j(\alpha_k \tau)) - \sigma(j_2)| \ge -c_7 \log p \tag{5.4}$$

for all $0 \le k \le p$ and all embeddings $\sigma : K(j(\alpha_0 \tau), \dots, j(\alpha_p \tau)) \to \mathbf{C}$ with $c_7 > 0$ independent of p, k, and σ.

Proof. All constants $c_3, \dots, c_6$ in this proof are positive and independent of p, k, and σ. Let $0 \le k \le p$. We note that $j(\alpha_k \tau)$ is a singular j-invariant because $\alpha_k \tau$ is imaginary quadratic. Since j_2 is not a singular j-invariant we have $j(\alpha_k \tau) \neq j_2$.

Let σ be as in the hypothesis. By Lemma 5.1 there is an imaginary quadratic $\rho \in \mathcal{F}$ such that $j(\rho) = j(\alpha_{\sigma(k)} \tau) = \sigma(j(\alpha_k \tau))$ with $h(\rho) \le c_2(1 + h(\alpha_{\sigma(k)} \tau))$. Moreover, by the first part of Lemma 5.1 we can estimate $h(\alpha_{\sigma(k)} \tau)$ from above by $c_1(1 + \log |\alpha_{\sigma(k)}| + h(\tau))$. Combining these estimates with $|\alpha_{\sigma(k)}| = p$ and using the fact that τ is fixed yields

$$h(\rho) \le c_3 \log p. \tag{5.5}$$

We write $\sigma(j_2) = j(\eta)$ with $\eta \in \mathcal{F}$. Here η depends on σ, but as j_2 is fixed, there are only finitely many possibilities for η. To prove the lemma we may assume that $|j(\rho) - j(\eta)|$ is small with respect to a constant depending on j_2. Because η and ρ are both in the fundamental domain we may assume that $|\rho - \eta|$ is small, after possibly replacing η by $\eta \pm 1$ or $-\eta^{-1}$ to deal with the usual boundary issues. We may even suppose that

$$|j(\rho) - j(\eta)| \ge \frac{1}{2} |j'(\eta)| |\rho - \eta|; \tag{5.6}$$

we note that $j'(\eta) \neq 0$ since η is not one of the three values $i, \exp(2\pi i/3), \exp(2\pi i/6)$ in $\mathcal{F}$ where j' vanishes. Hence $|j'(\eta)|$ is bounded from below by a positive constant depending only on j_2.

There is an elliptic curve defined over $\overline{\mathbf{Q}}$ with period lattice basis $\{\omega_1, \omega_2\} \subset \mathbf{C}$ such that $\eta = \omega_2/\omega_1$.

We consider the linear form $\omega_1 \rho - \omega_2$ in elliptic logarithms. If it were to vanish we would have $\rho = \eta$, hence $\sigma(j(\alpha_k \tau)) = j(\alpha_{\sigma(k)} \tau) = j(\rho) = j(\eta) = \sigma(j_2)$. But this is impossible by the first part of the lemma.

The coefficient ρ is an imaginary quadratic number; in particular, its degree is uniformly bounded. Moreover, $|\omega_1|$ and $|\omega_2|$ are bounded solely in terms of j_2.

The result of David and Hirata-Kohno gives

$$\log|\omega_1\rho-\omega_2|\geq -c_4(1+h(\rho)).$$

The estimate (5.5) implies $\log|\omega_1\rho-\omega_2|\geq -c_5\log p$. After dividing by ω_1 we obtain $\log|\rho-\eta|\geq -c_6\log p$. We recall (5.6) and conclude (5.4). □

5.2. **Equidistribution of Hecke orbits.** We continue to use notation introduced in the previous subsection.

For $z\in\mathbf{C}$ let $B_\delta(z)$ denote the open ball in $\mathbf{C}$ centered at z of radius δ. We define the open set

$$\mathcal{N}_\delta=j^{-1}\left(\bigcup_{\sigma:K\to\mathbf{C}}B_\delta(\sigma(j_2))\right)\subset\mathbf{H},$$

and for an embedding $\sigma:K\to\mathbf{C}$ introduce index sets

$$\mathcal{I}_\sigma(p,\delta)=\{0\leq k\leq p;\ \alpha_{\sigma(k)}\tau_\sigma\in\mathcal{N}_\delta\}.$$

The purpose of these index sets is to keep track of elements in the Hecke orbit of τ_σ whose image under j is near some $\sigma(j_2)$.

Clearly, $\#\mathcal{I}_\sigma(p,\delta)\leq p+1$. But for small δ one should be able to expect a better bound. The next lemma uses an equidistribution result of Clozel and Ullmo [7] and tells us just this.

Lemma 5.3. *If p is large enough then $\#\mathcal{I}_\sigma(p,\delta)\leq c_9\delta^2 p$ for all $\sigma:K\to\mathbf{C}$ with $c_9>0$ independent of p and δ.*

Proof. Let $\phi:\mathbf{H}\to\{0,1\}$ be the characteristic function of $\mathcal{N}_\delta$ and let $\tilde\phi$ be a continuous, real valued function on $\mathbf{H}$ with compact support such that $|\tilde\phi-\phi|\leq\delta^2$ on $\mathbf{H}$. By Clozel and Ullmo's Théorème 2.1(c) [7] we have

$$\frac{1}{p+1}\sum_{k=0}^{p}\tilde\phi(\alpha_{\sigma(k)}\tau_\sigma)\to\frac{3}{\pi}\int_{\mathcal{F}}\tilde\phi\frac{dxdy}{y^2}\quad\text{as}\quad p\to\infty$$

where x,y are the real and imaginary part functions on $\mathbf{H}$. The factor $3/\pi$ makes sure that $\mathcal{F}$ has total measure 1. If p is large enough we have

$$\frac{1}{p+1}\sum_{k=0}^{p}\tilde\phi(\alpha_{\sigma(k)}\tau_\sigma)\leq\frac{3}{\pi}\int_{\mathcal{F}}\tilde\phi\frac{dxdy}{y^2}+\delta^2.$$

The left-hand side above is at least $\frac{1}{p+1}\#\mathcal{I}_\sigma(p,\delta)-\delta^2$. The right-hand side can be bounded using

$$\frac{3}{\pi}\int_{\mathcal{F}}\tilde\phi\frac{dxdy}{y^2}\leq\frac{3}{\pi}\int_{\mathcal{F}}\phi\frac{dxdy}{y^2}+\delta^2\frac{3}{\pi}\int_{\mathcal{F}}\frac{dxdy}{y^2}=\frac{3}{\pi}\int_{\mathcal{F}}\phi\frac{dxdy}{y^2}+\delta^2.$$

The set $\bigcup_\sigma B_\delta(\sigma(j_2))$ has measure at most $[\mathbf{Q}(j_2):\mathbf{Q}]\pi\delta^2$. Elementary calculus gives

$$\frac{3}{\pi}\int_{\mathcal{F}}\phi\frac{dxdy}{y^2}\leq c_8\delta^2$$

with $c_8>0$ independent of p and δ.

Thus for p large enough we obtain $\#\mathcal{I}_\sigma(p,\delta) \leq (c_8+3)\delta^2(p+1) \leq 2(c_8+3)\delta^2 p$. □

The following lemma, whose proof rests on elementary methods, will be useful later on.

Lemma 5.4. *There exists $\epsilon(\delta) \in (0,1/2]$, which may depend on δ but not on p, with the following property: if p is an arbitrary prime then*

$$|\sigma(j(\alpha_k\tau)) - \sigma(j_2)| \geq \epsilon(\delta)\max\{1, |\sigma(j(\alpha_k\tau))|\}$$

for all $k \in \{0,\dots,p\}\backslash\mathcal{I}_\sigma(p,\delta)$ and all $\sigma : K(j(\alpha_0\tau),\dots,j(\alpha_p\tau)) \to \mathbf{Q}$. Moreover, $\log\max\{1,|\sigma(j(\alpha_k\tau))|\} \leq c_{10}$ for all $k \in \mathcal{I}_\sigma(p,\delta)$ and for all σ as before with $c_{10} > 0$ independent of p, δ, σ, and k.

Proof. Let σ be as in the hypothesis. The continuous function

$$q : \mathbf{H}\backslash\mathcal{N}_\delta \to [0,\infty) \quad \text{given by} \quad \rho \mapsto \frac{|j(\rho)-\sigma(j_2)|}{\max\{1,|j(\rho)|\}}$$

does not vanish by the definition of $\mathcal{N}_\delta$. Let us consider its restriction to $\mathcal{F}$: since $\mathcal{F}\backslash(\mathcal{F}\cap\mathcal{N}_\delta)$ is closed and since $q|_{\mathcal{F}}(\tau) \to 1$ as $\tau \to +\infty i$ we see that $q|_{\mathcal{F}}$ is uniformly bounded from below by some $\epsilon(\delta) \in (0,1/2]$. The set $\mathbf{H}\backslash\mathcal{N}_\delta$ and the function q are invariant under the action of $\mathrm{SL}_2(\mathbf{Z})$. The first part of the lemma follows since $\alpha_{\sigma(k)}\tau_\sigma \in \mathbf{H}\backslash\mathcal{N}_\delta$ for k as in the hypothesis.

To prove the second part we shall assume $k \in \mathcal{I}_\sigma(p,\delta)$. By definition there exists an embedding $\sigma' : K \to \mathbf{C}$ such that $|j(\alpha_{\sigma(k)}\tau_\sigma) - \sigma'(j_2)| < \delta$. Hence $|\sigma(j(\alpha_k\tau))| = |j(\alpha_{\sigma(k)}\tau_\sigma)| \leq \delta + |\sigma'(j_2)| \leq 1 + |\sigma'(j_2)|$ and the lemma follows. □

5.3. **Combining the estimates.** We continue to use the notation from the previous two subsections. We combine the estimates obtained therein to get the next lemma.

Lemma 5.5. *We have $\Phi_p(j_1,j_2) \neq 0$ and for p large enough*

$$\begin{aligned}(5.7)\qquad &\sum_{\sigma:K\to\mathbf{C}} \log|\sigma(\Phi_p(j_1,j_2))| \\ &\geq -c_{13}p(|\log\epsilon(\delta)| + \delta^2\log p) + \sum_{k=0}^{p}\sum_{\sigma:K\to\mathbf{C}} \log\max\{1,|\sigma(j(\alpha_k\tau))|\}\end{aligned}$$

with $c_{13} > 0$ independent of p and δ.

Proof. The non-vanishing of $\Phi_p(j_1,j_2)$ follows from the first part of Lemma 5.2 and (5.2).

Let $\sigma : K \to \mathbf{C}$ be an embedding. We recall (5.3) and split the following sum up:

$$\begin{aligned}(5.8)\qquad &\log|\sigma(\Phi_p(j_1,j_2))| \\ &= \sum_{k\notin\mathcal{I}_\sigma(p,\delta)} \log|\sigma(j(\alpha_k\tau)) - \sigma(j_2)| + \sum_{k\in\mathcal{I}_\sigma(p,\delta)} \log|\sigma(j(\alpha_k\tau)) - \sigma(j_2)|.\end{aligned}$$

We proceed by estimating each of these two sums separately.

For $k \notin \mathcal{I}_\sigma(p,\delta)$ we apply Lemma 5.4 to obtain

$$\begin{aligned}\log|\sigma(j(\alpha_k\tau)) - \sigma(j_2)| &\geq \log\epsilon(\delta) + \log\max\{1, |\sigma(j(\alpha_k\tau))|\} \\ &= -|\log\epsilon(\delta)| + \log\max\{1, |\sigma(j(\alpha_k\tau))|\}.\end{aligned}$$

For $k \in \mathcal{I}_\sigma(p,\delta)$ we apply Lemma 5.2 to obtain

$$\log|\sigma(j(\alpha_k\tau)) - \sigma(j_2)| \geq -c_7 \log p.$$

We insert these two estimates into (5.8) and use $p+1 \leq 2p$ to see

$$\begin{aligned}(5.9)\qquad &\log|\sigma(\Phi_p(j_1,j_2))| \\ &\geq -2p|\log\epsilon(\delta)| - c_7 \#\mathcal{I}_\sigma(p,\delta)\log p + \sum_{k\notin\mathcal{I}_\sigma(p,\delta)} \log\max\{1,|\sigma(j(\alpha_k\tau))|\}.\end{aligned}$$

Next we may rewrite

$$\begin{aligned}\sum_{k\notin\mathcal{I}_\sigma(p,\delta)} \log\max\{1,|\sigma(j(\alpha_k\tau))|\} = &\sum_{k=0}^{p} \log\max\{1,|\sigma(j(\alpha_k\tau))|\} + \\ &- \sum_{k\in\mathcal{I}_\sigma(p,\delta)} \log\max\{1,|\sigma(j(\alpha_k\tau))|\}.\end{aligned}$$

By the second assertion of Lemma 5.4 any term in the final sum is at most c_{10}. So,

$$\sum_{k\notin\mathcal{I}_\sigma(p,\delta)} \log\max\{1,|\sigma(j(\alpha_k\tau))|\} \geq \sum_{k=0}^{p} \log\max\{1,|\sigma(j(\alpha_k\tau))|\} - c_{10}\#\mathcal{I}_\sigma(p,\delta).$$

We insert the previous inequality in (5.9) to obtain

$$\log|\sigma(\Phi_p(j_1,j_2))| \geq -2p|\log\epsilon(\delta)| - c_{11}\#\mathcal{I}_\sigma(p,\delta)\log p + \sum_{k=0}^{p} \log\max\{1,|\sigma(j(\alpha_k\tau))|\}$$

with $c_{11} > 0$ independent of p and δ. By Lemma 5.3 we have $\#\mathcal{I}_\sigma(p,\delta) \leq c_9\delta^2 p$ for p large enough, so

$$\log|\sigma(\Phi_p(j_1,j_2))| \geq -2p|\log\epsilon(\delta)| - c_{12}\delta^2 p\log p + \sum_{k=0}^{p} \log\max\{1,|\sigma(j(\alpha_k\tau))|\}$$

with $c_{12} > 0$ independent of p and δ.

We take the sum over all embeddings $\sigma : K \to \mathbf{C}$ to conclude the proof. □

For an integer $N \geq 1$ the complex number $j(N\tau)$ is the j-invariant of an elliptic curve E_N. Then E_N has complex multiplication, in other words $j(N\tau)$ is a singular j-invariant, since $\Phi_N(j_1, j(N\tau)) = 0$. In the next lemma we rewrite the lower bound given in the previous lemma in terms of the height of the algebraic number $j(p\tau)$.

Let $L \subset \mathbf{C}$ be the imaginary quadratic number field generated by τ. For brevity, we write $\Delta = \Delta_L, \mathcal{O} = \mathcal{O}_L$, and $\chi = \chi_L$. For an integer f we write $\mathcal{O}_f = \mathbf{Z} + f\mathcal{O}$. Now $\mathrm{End}(E_1)$, being an order in $\mathcal{O}$, is of the form $\mathcal{O}_f$ for a

unique integer $f \geq 1$. There are coprime integers a, b, and c with $a \neq 0$ such that $a\tau^2 + b\tau + c = 0$. Lemma 7.5 [11] implies $\mathrm{End}(E_1) = \mathbf{Z} + a\tau\mathbf{Z}$. The same lemma also implies $\mathrm{End}(E_p) = \mathbf{Z} + pa\tau\mathbf{Z}$ provided p does not divide a. This assumption will turn out to be harmless since τ is fixed. If it holds true, then $\mathrm{End}(E_p) = \mathbf{Z} + fp\mathcal{O} = \mathcal{O}_{fp}$.

Lemma 5.6. *For p large enough we have*

$$\sum_{\sigma: K \to \mathbf{C}} \log |\sigma(\Phi_p(j_1, j_2))| \geq -c_{13} p(|\log \epsilon(\delta)| + \delta^2 \log p) + \frac{p}{4} h(j(p\tau))$$

with c_{13} from the previous lemma.

Proof. We consider the polynomial $A = \prod_\sigma \Phi_p(\sigma(j_1), Y)$ in one variable Y with rational coefficients. Now A vanishes at $j(p\tau)$, hence it vanishes at $\sigma(j(p\tau))$ for any embedding $\sigma : K(j(p\tau)) \to \mathbf{C}$. Since $j(p\tau)$ is an algebraic integer, our definition of the height (2.1) implies

$$[\mathbf{Q}(j(p\tau)) : \mathbf{Q}] h(j(p\tau)) \leq \sum_{\substack{z \in \mathbf{C} \\ A(z) = 0}} \log \max\{1, |z|\}.$$

By (5.3) the roots of A (counted with multiplicities) are in one-to-one correspondence with the terms in the sum on the right of (5.7). We obtain

(5.10)
$$\sum_{\sigma: K \to \mathbf{C}} \log |\sigma(\Phi_p(j_1, j_2))| \geq -c_{13} p(|\log \epsilon(\delta)| + \delta^2 \log p) + [\mathbf{Q}(j(p\tau)) : \mathbf{Q}] h(j(p\tau)).$$

We now bound $[\mathbf{Q}(j(p\tau)) : \mathbf{Q}]$ from below by using classical results from the theory of complex multiplication. The paragraph before this lemma implies $\mathrm{End}(E_p) = \mathcal{O}_{fp}$ if p is large enough. By Theorems 7.24 and 11.1 of [11] we have

$$[L(j(\tau)) : L] = \frac{h_L f}{[\mathcal{O}^\times : \mathcal{O}_f^\times]} \prod_{l \mid f} \left(1 - \chi(l) l^{-1}\right) \tag{5.11}$$

and

$$[L(j(p\tau)) : L] = \frac{h_L fp}{[\mathcal{O}^\times : \mathcal{O}_{fp}^\times]} \prod_{l \mid fp} \left(1 - \chi(l) l^{-1}\right),$$

the product runs over all primes l dividing f and fp, respectively. Certainly, f is fixed by τ so we may assume that p does not divide f. If $p \geq 5$ we obtain

$$\frac{[L(j(p\tau)) : L]}{[L(j(\tau)) : L]} = \frac{[\mathcal{O}^\times : \mathcal{O}_f^\times]}{[\mathcal{O}^\times : \mathcal{O}_{fp}^\times]} p \left(1 - \chi(p) p^{-1}\right) \geq \frac{p-1}{[\mathcal{O}^\times : \mathcal{O}_{fp}^\times]} \geq \frac{p-1}{3} \geq \frac{p}{4}$$

on using $[\mathcal{O}^\times : \mathcal{O}_{fp}^\times] \leq \#\mathcal{O}^\times/2 \leq 3$. We note $[\mathbf{Q}(j(p\tau)) : \mathbf{Q}] \geq [L(j(p\tau)) : L]$ (actually, equality holds). The lemma follows from (5.10) and $[L(j(p\tau)) : L] \geq p[L(j(\tau)) : L]/4 \geq p/4$. □

To get a lower bound for $h(j(p\tau))$ in terms of p we shall work with the Faltings height of the elliptic curve associated to $j(p\tau)$. We then apply the estimates of section 4.

If $j_E \in \overline{\mathbf{Q}}$ is the j-invariant of an elliptic curve E, we remind the reader that the second bound in (3.1) gives

$$h_F(E) \leq \frac{1}{12} h(j_E) + c_{14} \tag{5.12}$$

for some absolute constant c_{14}.

In the following lemma we make use of Lemma 4.2. It enables us to express $h_F(E_p)$ in terms of $h_F(E_1)$ and p.

Lemma 5.7. *For p large enough we have*

$$\sum_{\sigma : K \to \mathbf{C}} \log |\sigma(\Phi_p(j_1, j_2))| \geq -c_{15} p(|\log \epsilon(\delta)| + \delta^2 \log p) + \frac{6}{5} p \log p$$

with $c_{15} > 0$ independent of p and δ.

Proof. As we have seen above, $\mathrm{End}(E_1) = \mathbf{Z} + f\mathcal{O}$ and $\mathrm{End}(E_p) = \mathbf{Z} + fp\mathcal{O}$, at least if p is large enough.

On applying Lemma 4.2 we get the lower bound

$$h_F(E_p) \geq \frac{2}{5} \log p \tag{5.13}$$

for p large enough.

Lemma 5.6 combined with height comparison estimate (5.12) gives

$$\sum_{\sigma : K \to \mathbf{C}} \log |\sigma(\Phi_p(j_1, j_2))| \geq -c_{13} p(|\log \epsilon(\delta)| + \delta^2 \log p) + 3p h_F(E_p) - 3p c_{14}.$$

By (5.13) the height of E_p grows logarithmically in p, hence

$$\sum_{\sigma : K \to \mathbf{C}} \log |\sigma(\Phi_p(j_1, j_2))| \geq -c_{13} p(|\log \epsilon(\delta)| + \delta^2 \log p) - 3p c_{14} + \frac{6}{5} p \log p.$$

By construction $|\log \epsilon(\delta)| \geq \log 2$ so $-3pc_{14}$ is absorbed by the term to its left after replacing c_{13} by a larger constant still independent of p and δ. □

The lemma below is now a simple consequence of the lemma above. Its proof involves choosing an appropriate δ.

Lemma 5.8. *For p large enough we have*

$$\sum_{\sigma : K \to \mathbf{C}} \log |\sigma(\Phi_p(j_1, j_2))| \geq p \log p.$$

Proof. By Lemma 5.7 we have

$$\sum_{\sigma : K \to \mathbf{C}} \log |\sigma(\Phi_p(j_1, j_2))| \geq -c_{15} p(|\log \epsilon(\delta)| + \delta^2 \log p) + \frac{6}{5} p \log p$$

for p large enough. We recall that $c_{15} > 0$ is independent of δ. Hence we may fix $\delta \in (0, 1]$ such that $c_{15}\delta^2 < 1/10$. We obtain

$$\sum_{\sigma: K \to \mathbf{C}} \log |\sigma(\Phi_p(j_1, j_2))| \geq -c_{15} p |\log \epsilon(\delta)| + \frac{11}{10} p \log p.$$

The proposition follows since $\epsilon(\delta)$ is independent of p. □

5.4. **An application to curves in $\mathbf{A}^2$.** In this subsection we prove Theorem 1.2.

Let $F \in \mathbf{Z}[X, Y]$ be an irreducible and non-constant polynomial. We assume that it does not divide Φ_p for any prime p.

We keep much of the same notation as in the previous subsections with the important difference that we need a stronger hypothesis on j_1.

Let j_1 be a singular j-invariant such that

(5.14) if $j_2^* \in \overline{\mathbf{Q}}$ satisfies $F(j_1, j_2^*) = 0$ then j_2^* is not a singular j-invariant.

Finally, let j_2 be algebraic with $F(j_1, j_2) = 0$ and let K be a number field containing j_1 and j_2.

By hypothesis j_2 cannot be a singular j-invariant.

Our convention on constants stays the same except that $c_{16}, \ldots$ are now also allowed to depend on F.

Let p be a prime. In this subsection we call p large enough if it is larger than some constant which may depend on j_1, j_2, K, and F.

The polynomials F and Φ_p only have finitely many common zeros and these are algebraic. Let $x_1, \ldots, x_n$ be the distinct algebraic numbers which appear as a first coordinate of such a zero; here $n = 0$ is possible. Let R be the resultant of F and Φ_p considered as polynomials with coefficients in $\mathbf{Z}[X]$, cf. Ch. IV, §6, No. 6 of [5]. We remark that Φ_p is monic when considered as a polynomial in Y. Then

$$R = \lambda \prod_{i=1}^{n} (X - x_i)^{a_i} \in \mathbf{Z}[X]$$

for integers $a_i \geq 1$ and an integer $\lambda \neq 0$. Moreover, there exist polynomials $U, V \in \mathbf{Z}[X, Y]$ with

(5.15) $\deg_Y U < \deg_Y F$ and $\deg_Y V < \deg_Y \Phi_p$ such that $U\Phi_p + VF = R$.

We may assume that U, V, and R have no common integral prime factor. Our resultant is the determinant of a certain matrix with entries determined by the coefficients of Φ_p and F taken as a polynomial in one variable Y. A straightforward degree estimate applied to this determinant shows

(5.16)

$$\deg R \leq \deg_X \Phi_p \deg_Y F + \deg_Y \Phi_p \deg_X F = (p+1)(\deg_X F + \deg_Y F) \leq 4p \deg F.$$

It is important to keep in mind that that U, V, and R mostly likely depend on the prime p.

We recall that a polynomial in integer coefficients is called primitive if its coefficients have no common prime divisor.

Lemma 5.9. *If $p > \deg F$, then R is primitive.*

Proof. Let l be a prime. If A is any polynomial in integer coefficient we let $\overline{A}$ denote its reduction modulo l.

We assume that $p > \deg F$ and $\overline{R} = 0$ and shall derive a contradiction.

The resultant equation (5.15) modulo l reads

$$\overline{U}\,\overline{\Phi_p} + \overline{V}\,\overline{F} = 0.$$

We begin by remarking that $\overline{F} \neq 0$ because F is irreducible in $\mathbf{Z}[X,Y]$ and non-constant. The leading term of Φ_p taken as a polynomial with coefficients in $\mathbf{Z}[X]$ is Y^{p+1}, thus $\overline{\Phi_p} \neq 0$ and even $\deg_Y \overline{\Phi_p} = \deg_Y \Phi_p = p+1$. If $\overline{V} = 0$ then we would have $\overline{U} = 0$, but this is impossible because l cannot be a common factor of U, V, and R. Hence $\overline{V} \neq 0$.

Let us assume $l \neq p$ for the moment. By a result of Igusa [18] the polynomial $\overline{\Phi_p}$ is absolutely irreducible. Hence it must divide $\overline{V}$ or $\overline{F}$. The former situation is impossible since $\overline{V} \neq 0$ and $\deg_Y \overline{V} \leq \deg_Y V < \deg_Y \Phi_p = \deg_Y \overline{\Phi_p}$, here we used (5.15). The latter situation cannot hold either since $\overline{F} \neq 0$ and $\deg_Y \overline{F} \leq \deg F < p < \deg_Y \overline{\Phi_p}$.

Therefore, we must have $l = p$. In this case $\overline{\Phi_p}$ is no longer irreducible since Kronecker's congruence relation, §2 Chapter 5 of [19], implies $\overline{\Phi_p} = (X^p - Y)(X - Y^p)$. But both factors $X^p - Y$ and $X - Y^p$ are (absolutely) irreducible. Because $\deg_X(X^p - Y) = p > \deg F \geq \deg_X \overline{F}$ and $\deg_Y(X - Y^p) = p > \deg F \geq \deg_Y \overline{F}$ we deduce that $X^p - Y$ and $X - Y^p$ both divide the non-zero $\overline{V}$. This conclusion implies $\deg_Y \overline{V} \geq \deg_Y \overline{\Phi_p}$ and so $\deg_Y V \geq \deg_Y \Phi_p$. By (5.15) we have arrived at a contradiction. □

If $U(j_1, j_2)$ is non-zero we can use the product formula to obtain a lower bound for $|R|$, the largest absolute value of all coefficients of R.

Lemma 5.10. *If $U(j_1, j_2) \neq 0$, then*

$$\log |R| \geq -c_{16} p + \frac{1}{[K : \mathbf{Q}]} \sum_{\sigma : K \to \mathbf{C}} \log |\sigma(\Phi_p(j_1, j_2))|,$$

with $c_{16} > 0$ independent of p.

Proof. If $v \in M_K^0$ then $|j_1|_v \leq 1$ because j_1 is an algebraic integer, being a singular j-invariant. Since U has integer coefficients and by the ultrametric triangle inequality we deduce $|U(j_1, j_2)|_v \leq \max\{1, |j_2|_v\}^{\deg_Y U}$. By (5.15) the degree $\deg_Y U$ is at most $\deg_Y F$, so

$$|U(j_1, j_2)|_v \leq \max\{1, |j_2|_v\}^{\deg_Y F}. \tag{5.17}$$

By (5.15) and because of $F(j_1, j_2) = 0$ we get

$$\prod_{\sigma : K \to \mathbf{C}} |\sigma(R(j_1))| = \prod_{\sigma : K \to \mathbf{C}} |\sigma(U(j_1, j_2))||\sigma(\Phi_p(j_1, j_2))|. \tag{5.18}$$

Since $U(j_1, j_2) \neq 0$ we have the product formula:

$$1 = \prod_{\sigma: K \to \mathbf{C}} |\sigma(U(j_1, j_2))| \prod_{v \in M_K^0} |U(j_1, j_2)|_v^{d_v}.$$

We estimate the product over M_K^0 from above by using (5.17) to obtain

$$1 \leq \prod_{\sigma: K \to \mathbf{C}} |\sigma(U(j_1, j_2))| \prod_{v \in M_K^0} \max\{1, |j_2|_v\}^{d_v \deg_Y F}.$$

Multiplying both sides with $\prod_\sigma |\sigma(\Phi_p(j_1, j_2))|$ and applying (5.18) gives

$$\prod_{\sigma: K \to \mathbf{C}} |\sigma(\Phi_p(j_1, j_2))| \leq \prod_{\sigma: K \to \mathbf{C}} |\sigma(R(j_1))| \prod_{v \in M_K^0} \max\{1, |j_2|_v\}^{d_v \deg_Y F}. \tag{5.19}$$

The triangle inequality implies

$$|\sigma(R(j_1))| = |R(\sigma(j_1))| \leq (1 + \deg R)|R| \max\{1, |\sigma(j_1)|\}^{\deg R}$$

for any $\sigma : K \to \mathbf{C}$. We recall (5.16) to deduce

$$|\sigma(R(j_1))| \leq 8p \deg(F)|R| \max\{1, |\sigma(j_1)|\}^{4p \deg F}.$$

Inserting this into (5.19) and taking the $[K : \mathbf{Q}]$th root gives

$$\prod_{\sigma: K \to \mathbf{C}} |\sigma(\Phi_p(j_1, j_2))|^{1/[K:\mathbf{Q}]} \leq$$
$$8p \deg(F)|R| \prod_{\sigma: K \to \mathbf{C}} \max\{1, |\sigma(j_1)|\}^{\frac{4p \deg F}{[K:\mathbf{Q}]}} \prod_{v \in M_K^0} \max\{1, |j_2|_v\}^{\frac{d_v \deg_Y F}{[K:\mathbf{Q}]}}.$$

because there are precisely $[K : \mathbf{Q}]$ embeddings σ. Using the definition of the height given in section 2 we obtain

$$\prod_{\sigma: K \to \mathbf{C}} |\sigma(\Phi_p(j_1, j_2))|^{1/[K:\mathbf{Q}]} \leq 8p \deg(F)|R| H(j_1)^{4p \deg F} H(j_2)^{\deg_Y F}.$$

The lemma follows from an elementary calculation. □

Lemma 5.11. *For p large enough there exists a common zero (x, y) of Φ_p and F such that*

$$h(x, y) \geq c_{18} \log p,$$

with $c_{18} > 0$ independent of p.

Proof. Let us assume for the moment that $U(j_1, j_2) \neq 0$. We apply logarithm to the estimate in Lemma 5.10 and use the lower bound for $\sum_\sigma \log |\sigma(\Phi_p(j_1, j_2))|$ from Lemma 5.8 to deduce

$$\log |R| \geq -c_{16} p + \frac{1}{[K : \mathbf{Q}]} p \log p \geq c_{17} p \log p \tag{5.20}$$

for p large enough and with $c_{17} > 0$ independent of p.

We may certainly assume that $p > \deg F$. Lemma 5.9 implies that R is primitive. If R were a constant, in other words if Φ_p and F had no common zeros, we

would have $R = \pm 1$. By (5.20) this is impossible for p large enough. So Φ_p and F will eventually have common zeros as p increases.

But we get more information. Indeed, the polynomial R factors over $\mathbf{Z}[X]$ into $\pm R_1^{r_1} \cdots R_g^{r_g}$ where $R_i \in \mathbf{Z}[X]$ are irreducible and $r_i \geq 1$. Since R is primitive, the R_i are non-constant. By the previous paragraph we may assume $g \geq 1$.

Let $m(\cdot)$ denote the (logarithmic) Mahler measure of a polynomial with complex coefficients. After rearranging the $R_i's$ we may suppose $m(R_1)/\deg R_1 \geq m(R_i)/\deg R_i$ for $1 \leq i \leq g$. The Mahler measure is additive, so

$$m(R) = \sum_{i=1}^{g} r_i m(R_i) \leq \left(\sum_{i=1}^{g} r_i \deg R_i \right) \frac{m(R_1)}{\deg R_1} = \deg R \frac{m(R_1)}{\deg R_1}.$$

By Lemma 1.6.7 of [2] we can bound $\log |R| \leq \deg R \log 2 + m(R)$. We use (5.16) to bound $\deg R$ and deduce

$$\log |R| \leq \deg R \left(\log 2 + \frac{m(R_1)}{\deg R_1} \right) \leq 4p \deg F \left(\log 2 + \frac{m(R_1)}{\deg R_1} \right).$$

We combine this inequality with the lower bound (5.20) and cancel p. Fortunately, the lower bound for $\log |R|$ also contains a factor $\log p$; we get

$$\frac{m(R_1)}{\deg R_1} \geq c_{18} \log p$$

for p large enough.

By Gauss's Lemma, R_1 is primitive. Since R_1 is non-constant it has a root x. Proposition 1.6.6 of [2] and $\deg R_1 = [\mathbf{Q}(x) : \mathbf{Q}]$ imply the equality in

$$h(x) = \frac{m(R_1)}{\deg R_1} \geq c_{18} \log p.$$

Because $R(x) = 0$ and by properties of the resultant mentioned earlier, there exists an algebraic y such that (x, y) is a common zero of Φ_p and F. Certainly, $h(x, y) \geq h(x)$ and so the lemma holds in the case $U(j_1, j_2) \neq 0$.

But what if $U(j_1, j_2) = 0$? This looks troubling since U was constructed using Φ_p and will most likely depend on p. However, if $U(j_1, j_2) = 0$ for some p, then (5.15) implies $R(j_1) = 0$. By properties of the resultant, j_1 is the first coordinate of a common zero of F and Φ_p. In other words, there exists j_2^* such that $\Phi_p(j_1, j_2^*) = F(j_1, j_2^*) = 0$. Now j_2^* must be a singular j-invariant, a contradiction in view of our hypothesis on j_1 given in (5.14). So, $U(j_1, j_2)$ cannot vanish for any p. □

5.5. **Proof of Theorem 1.2.** Let C be the curve from the hypothesis. After interchanging coordinates we may assume that C is not a vertical line. So its projection onto the first factor of $\mathbf{A}^2$ contains all but finitely many points. Let $C_1, \ldots, C_r$ be the conjugates over $\mathbf{Q}$ of C. None of the C_i is a special curve by hypothesis.

The André-Oort Conjecture is known unconditionally for curves in $\mathbf{A}^2$ by a result of André [1]. Therefore, the set of points in $C_1(\overline{\mathbf{Q}}) \cup \cdots \cup C_r(\overline{\mathbf{Q}})$ whose coordinates are both singular j-invariants, is finite. Among the infinitely many

singular j-invariants we can choose one, say j_1, that does not appear as the first coordinate of a point in this set. We may even choose j_1 such that it is contained in the projection of C onto the first factor of $\mathbf{A}^2$. So $(j_1, j_2) \in C(\overline{\mathbf{Q}})$ for some algebraic j_2.

The curve $C_1 \cup \cdots \cup C_r$ is defined and irreducible over $\mathbf{Q}$. Hence it is the zero set of a single non-constant and irreducible polynomial $F \in \mathbf{Z}[X, Y]$. Certainly, F does not divide any Φ_p. Our choice of j_1 satisfies the condition (5.14) with respect to this polynomial.

We apply Lemma 5.11 to F and the auxiliary point (j_1, j_2). As a consequence there is $c > 0$ such that for p large enough we may find a common zero of Φ_p and F with height at least $c \log p$. By construction, some conjugate of this zero lies in $C(\overline{\mathbf{Q}})$. It remains a zero of Φ_p. Our theorem follows because the height is invariant under conjugation. □

The values $j_1 = j((\sqrt{-7}+1)/2) = -3375$ and $j_2 = -3375 - 1$ determine an admissible auxiliary point for the polynomial $F = X - Y - 1$. Indeed, j_1 is a singular j-invariant and j_2 is not on the well-known list of rational singular j-invariants. So $K = \mathbf{Q}$ is possible here. By going through the proofs of this section and optimizing the estimates for this special case, one can show the following: if C is the curve defined by F then for all large enough primes p there is $P \in (C \cap Y_0(p))(\overline{\mathbf{Q}})$ with $h(P) \geq 3 \log p - c \log \log p$, here c is an absolute constant. It follows that the constant 6 in Theorem 1.3 cannot be replaced by anything strictly less than 3.

6. The curve $X - Y = 1$ and proof of Theorem 1.3

Let p be a prime and let us define

$$R = -\Phi_p(X, X - 1) \in \mathbf{Z}[X].$$

The following simple observation is crucial for obtaining the height bound in Theorem 1.3: by Kronecker's congruence relation the reduction of Φ_p modulo p is $(X^p - Y)(X - Y^p)$, so the reduction of R modulo p is

$$(X^p - X + 1)(X^p - X - 1).$$

The factors $X^p - X + 1$ and $X^p - X - 1$ are Artin-Schreier polynomials and they are irreducible over $\mathbf{F}_p$. We will see in a moment that R is monic of degree $2p$. There are now two possibilities for R: it is either irreducible in $\mathbf{Z}[X]$ or factors into two irreducible polynomials in $\mathbf{Z}[X]$ of degree p.

We now show that R is monic of degree $2p$. If p is odd the Kronecker congruence relation and Theorem 3(iii), Chapter 5 of [19], imply that $\Phi_p(X, X) \in \mathbf{Z}[X]$ has degree $2p$ and leading coefficients -1; there is a misprint in the statement of Theorem 3(iii): the proof gives leading coefficient ± 1. Since the explicit form of Φ_2 is known, one directly checks that $\Phi_2(X, X)$ has degree 4 and leading coefficient -1. If $\Phi_p = \sum_{i,j} a_{ij} X^i Y^j$ we must have $\sum_{i+j=2p} a_{ij} = -1$. We recall that Φ_p is monic of degree $p + 1$ as a polynomial with coefficients in $\mathbf{Z}[X]$. The same holds true when considered as a polynomial with coefficients in $\mathbf{Z}[Y]$. From this we can quickly deduce that $\deg R \leq 2p$. Moreover, if $a_{ij} \neq 0$ with

$i+j=2p$, then $i=j=p$. Thus $a_{pp}=-1$. This is the leading coefficient of $\Phi_p(X,X-1)=-R$, so R is monic of degree $2p$.

Any x with $(x,x-1)\in Y_0(p)$ satisfies $R(x)=0$, so it is algebraic. If R is irreducible over $\mathbf{Z}$, then by Proposition 1.6.6 [2] we have $h(x)=m(R)/\deg R=m(R)/(2p)$. In the other case $R=AB$ factors with $A,B\in\mathbf{Z}[X]$ irreducible of degree p. Say $A(x)=0$, then $h(x)=m(A)/\deg A=m(A)/p=m(R)/p-m(B)/p\leq m(R)/p$ since the Mahler measure is additive and non-negative. In any case we have

$$h(x)\leq\frac{m(R)}{p}\quad\text{and}\quad[\mathbf{Q}(x):\mathbf{Q}]\geq p. \tag{6.1}$$

This degree lower bound justifies the comment made in the introduction after the statement of the current theorem.

We continue by bounding $m(R)$ from above. If the a_{ij} are as above, then

$$R=\sum_{\substack{0\leq i,j\leq p+1\\0\leq k\leq j}}\pm a_{ij}\binom{j}{k}X^{i+k}=\sum_{d=0}^{2p}\left(\sum_{\substack{0\leq i,j\leq p+1\\i+k=d}}\pm a_{ij}\binom{j}{k}\right)X^d.$$

Elementary estimates give

$$\begin{aligned}|R|&\leq\max_{0\leq d\leq 2p}\sum_{\substack{0\leq i,j\leq p+1\\i+k=d}}|a_{ij}|\binom{j}{k}\\&\leq|\Phi_p|\max_{0\leq d\leq 2p}\sum_{\substack{0\leq i,j\leq p+1\\i+k=d}}\binom{j}{k}\leq|\Phi_p|\sum_{j=0}^{p+1}2^j\leq|\Phi_p|2^{p+2}.\end{aligned}$$

We apply Lemma 1.6.7 of [2] to get $m(R)\leq\frac{1}{2}\log(2p+1)+\log|R|$. With the estimate above we conclude

$$m(R)\leq\log|\Phi_p|+c_1p, \tag{6.2}$$

here and below c_1,c_2, and c_3 denote positive absolute constants.

Cohen's estimate [9] for the modular transformation polynomials provides $\log|\Phi_p|\leq 6p\log p+c_2p$. Combining this with (6.1) and (6.2) leads to $h(x)\leq 6\log p+c_3$.

Finally, the theorem follows because $h(x,y)=\max\{h(x),h(y)\}$ and $h(y)=h(x-1)\leq\log 2+h(x)$ by elementary height properties. □

What happens for the curve $X+Y=1$? The reduction of $\Phi_p(X,1-X)$ modulo p is

$$(X^p+X-1)^2$$

by Kronecker's congruence relation. In characteristic $p>2$ the factor X^p+X-1 has a root at $1/2$. Worse still, there is $\zeta\in\mathbf{F}_{p^2}$ with $\zeta^{p-1}=-1$. One easily verifies, that $1/2,1/2+\zeta,\dots,1/2+(p-1)\zeta$ are roots of X^p+X-1. This polynomial factors over $\mathbf{F}_p$ into one linear term and $(p-1)/2$ quadratic terms.

7. Proof of Corollary 1.2

Let C be as in the hypothesis. We begin by showing part (i).

Let $P \in C(\overline{\mathbf{Q}}) \cap \mathcal{S}'$. If one coordinate is a singular j-invariant, then so is the other by the definition of $\mathcal{S}'$. By the André-Oort Conjecture and after discarding finitely many points we may assume that neither coordinate of P is a singular j-invariant.

So $\mathcal{S}_{\mathbf{Q}}(P) = Y_0(N)$ for some $N \geq 1$; hence $\deg \mathcal{S}_{\mathbf{Q}}(P) = 2\Psi(N)$ with $\Psi(N)$ the arithmetic function mentioned in the introduction. Theorem 1.1 implies $h(P) \leq c \log(1+N)$ with c independent of N. Now $\Psi(N) \geq N$ easily implies this part of the corollary.

We now turn to part (ii). Let $P \in C(\overline{\mathbf{Q}}) \cap \mathcal{S}$. By the first part, we may assume $P \notin \mathcal{S}'$. We are in one of the first three cases of (1.3).

Let us assume $\mathcal{S}(P) = \{P\}$ or $\mathcal{S}(P) = \{x\} \times \mathbf{A}^1$. The first coordinate x of P is a singular j-invariant and (1.4) tells us

$$\deg \mathcal{S}_{\mathbf{Q}}(P) \geq [\mathbf{Q}(x) : \mathbf{Q}].$$

Let E be the elliptic curve with j-invariant x and let L denote its field of complex multiplication. We note $|\Delta_L| \geq 3$. Let f be the unique positive integer with $\mathrm{End}(E) = \mathbf{Z} + f\mathcal{O}$ where $\mathcal{O} = \mathcal{O}_L$. We shall assume, as we may, that $|\Delta_L| f^2$ is large enough. By applying (5.11) with $j(\tau)$ replaced by x we see $[L(x) : L] \geq h_L f(\prod_{l|f} 1 - l^{-1})/3$; here we used $[\mathcal{O}^\times : (\mathbf{Z} + f\mathcal{O})^\times] \leq \#\mathcal{O}^\times/2 \leq 3$. The expression $f \prod_{l|f}(1 - l^{-1})$ is Euler's quotient function evaluated at f and so at least $c_1 f^{1/2}$ with $c_1 > 0$ absolute. We conclude $\deg \mathcal{S}_{\mathbf{Q}}(P) = [\mathbf{Q}(x) : \mathbf{Q}] \geq [L(x) : L] \geq c_1 h_L f^{1/2}/3$.

For any $\epsilon > 0$, the Theorem of Brauer-Siegel implies $h_L \geq c_2 |\Delta_L|^{1/2-\epsilon}$, here $c_2 > 0$ depends only on ϵ. Picking $\epsilon = 1/4$ and using the conclusion of previous paragraph gives

$$\log(1 + \deg \mathcal{S}_{\mathbf{Q}}(P)) \geq c_3 \log(|\Delta_L| f^2) \tag{7.1}$$

with $c_3 > 0$ absolute.

Recall that we are assuming the GRH. By Lemmas 4.1 and 4.3 we can estimate $h_F(E) \leq c_4 \log(|\Delta_L| f^2)$ with $c_4 > 0$ absolute. The height $h(x)$ can be bounded above in terms of the Faltings height $h_F(E)$ by (3.1). We conclude $h(x) \leq c_5 \log(|\Delta_L| f^2)$ with $c_5 > 0$ absolute. The lower bound (7.1) implies

$$h(x) \leq c_6 \log(1 + \deg \mathcal{S}_{\mathbf{Q}}(P)) \tag{7.2}$$

with $c_6 > 0$ absolute.

The two heights $h(x)$ and $h(y)$ are not unrelated. If d_1 and d_2 are the degrees of the first, respectively second coordinate function on C we have $|h(x)d_2 - h(y)d_1| \leq c_7(1 + h(x))^{1/2}$ by (3.2) with $c_7 > 0$ independent of x and y. If $d_1 \neq 0$ we can bound $h(y)$ linearly from above in $h(x)$. The current case follows from (7.2).

It remains to show that d_1 cannot vanish. Indeed, if $d_1 = 0$ we would have $C = \{x\} \times \mathbf{A}^1$. But x is singular j-invariant. Then C would be a special curve in contradiction to our hypothesis.

The final case $\mathcal{S}(P) = \mathbf{A}^1 \times \{y\}$ with y a singular j-invariant can be handled in a similar manner.

Acknowledgments

I would like to thank Lars Kühne and Gisbert Wüstholz for fruitful discussions especially in connection with the former's Master Thesis [20] on the André-Oort Conjecture.

References

[1] Y. André, Finitude des couples d'invariants modulaires singuliers sur une courbe algébrique plane non modulaire, *J. Reine Angew. Math.* **505** (1998), 203–208.
[2] E. Bombieri and W. Gubler, *Heights in Diophantine Geometry*, Cambridge University Press, 2006.
[3] E. Bombieri, D. Masser, and U. Zannier, Intersecting a curve with algebraic subgroups of multiplicative groups, *Internat. Math. Res. Notices* no. 20 (1999), 1119–1140.
[4] E. Bombieri, D. Masser, and U. Zannier, Anomalous subvarieties - structure theorems and applications, *Int. Math. Res. Not. IMRN* no. 19 (2007), 1–33.
[5] N. Bourbaki, *Éléments de Mathématique. Algèbre*, ch. 4 à 7, Masson, Paris, 1981.
[6] J. W. S. Cassels, *An Introduction to Diophantine Approximation*, Cambridge University Press, 1957.
[7] L. Clozel and E. Ullmo, *Équidistribution des points de Hecke*, Contributions to automorphic forms, geometry, and number theory (Haruzo Hida et. al., ed.), Johns Hopkins Univ. Press, 2004, pp. 193–254.
[8] H. Cohen, *Number Theory Volume II: Analytic and Modern Tools*, Springer, 2007.
[9] P. Cohen, On the coefficients of the transformation polynomials for the elliptic modular function, *Math. Proc. Camb. Phil. Soc.* **95** (3) (1984), 389–402.
[10] P. Colmez, Sur la hauteur de Faltings des variétés abéliennes à multiplication complexe, *Compositio Math.* **111** (3) (1998), 359–368.
[11] D. A. Cox, *Primes of the form* $x^2 + ny^2$, John Wiley & Sons, 1989.
[12] H. Davenport, *Multiplicative Number Theory*, Springer, 2000.
[13] S. David and N. Hirata-Kohno, Linear forms in elliptic logarithms, *J. Reine Angew. Math.* **628** (2009), 37–89.
[14] P. Deligne, Preuve des conjectures de Tate et de Shafarevitch, Séminaire Bourbaki 1983/84, *Astérisque* **121–122** (1985), 25–41.
[15] B. Edixhoven, Special points on products of modular curves, *Duke Math. J.* **126** (2) (2005), 325–348.
[16] G. Faltings, Endlichkeitssätze für abelsche Varietäten über Zahlkörpern, *Invent. Math.* **73** (1983), 349–366.
[17] A. Granville and H. M. Stark, ABC implies no "Siegel zeros" for L-functions of characters with negative discriminant, *Invent. Math.* **139** (2000), 509–523.
[18] J. Igusa, Fibre systems of Jacobian varieties. III. Fibre systems of elliptic curves, *Amer. J. Math.* **81** (1959), 453–476.
[19] S. Lang, *Elliptic Functions*, Springer, 1987.
[20] L. Kühne, *The André-Oort Conjecture in* $\mathbf{C}^2$, Master Thesis ETH Zurich, 2009.
[21] D. W. Masser and G. Wüstholz, Estimating isogenies on elliptic curves, *Invent. Math.* **100** (1990), 1–24.
[22] G. Maurin, Courbes algébriques et équations multiplicatives, *Math. Ann.* **341** (2008), 789–824.
[23] Y. Nakkajima and Y. Taguchi, A generalization of the Chowla-Selberg formula, *J. Reine Angew. Math.* **419** (1991), 119–124.
[24] A. Néron, Quasi-fonctions et hauteurs sur les variétés abéliennes, *Ann. of Math.* **82** (2) (1965), 249–331.

[25] R. Pink, *A common generalization of the conjectures of André-Oort, Manin-Mumford, and Mordell-Lang*, Preprint (2005), 13pp.
[26] M. Raynaud, *Hauteurs et isogénies*, Seminar on arithmetic bundles: the Mordell conjecture (Paris, 1983/84). Astérisque, no. 127, 1985, pp. 199–234.
[27] J. H. Silverman, *Heights and Elliptic Curves*, Arithmetic Geometry (Gary Cornell and Joseph H. Silverman, eds.), Springer, 1986, pp. 253–265.

ETH Zurich, Raemistrasse 101
8092 Zurich, Switzerland
E-mail address: habegger@math.ethz.ch