Bulletin of the Institute of Mathematics
Academia Sinica (New Series)
Vol. **9** (2014), No. 4, pp. 735-779

A NOTE ON THE CONJECTURES OF ANDRÉ–OORT AND PINK

WITH AN APPENDIX BY LARS KÜHNE

GISBERT WÜSTHOLZ

Departement Mathematik, ETH-Zentrum Rämistrasse 101, CH-8092 Zürich & Institut für Analysis und Computational Number Theory (Math A), Technische Universität Graz, Steyrergasse 30 - A-8010.
E-mail: wustholz@math.ethz.ch

Abstract

This note consists of two parts. In the first we give a – as we believe – more conceptual proof of a slightly sharper effective version of a very nice result published by Kühne on the André-Oort conjecture for curves in $\mathbb{A}^1 \times \mathbb{A}^1$. The second part deals with an extension of the André-Oort conjecture by Pink where Shimura varieties are replaced by mixed Shimura varieties. We consider the particular case when the mixed Shimura variety is the product of two universal elliptic curves.

1. Introduction

In 1989 Y. André [1] and independently F. Oort [33] in 1997 published conjectures which intimately relate arithmetic and geometry on Shimura varieties and which led to a conjecture now commonly called the André-Oort conjecture. The first case of the conjecture was verified by André [2] himself in 1998 by using diophantine approximations and class field theory. His method resembles the method used by Siegel in his proof of the finiteness of integral points on curves. In the last 15 years there was much activity and progress towards a proof of the conjecture and some further particular cases have been verified including statements that the generalized Riemann hypothesis implies the conjecture. For an overview of the state of art see Scanlon's Bourbaki article [37] and also [32].

Received January 28, 2014 and in revised form August 26, 2014.

AMS Subject Classification: 11J86, 11J89, 11G18, 14G05, 14K20, 14K22.

Key words and phrases: André - Oort conjecture, Pink conjecture, linear forms in logarithms, mixed Shimura varieties, variation of Faltings height.

It remained however an open question whether an effective version of André 's result or more generally an effective proof of the general conjecture can be obtaind. The first step towards this general problem was done by Kühne in [23]. In his paper Kühne considers - as André does - a geometrically irreducible curve $\mathcal{C}_K$ in $\mathbb{A}^1 \times \mathbb{A}^1$ defined over a number field K and of bi-degree (d_1, d_2) with $d_1 d_2 \neq 0$ and with height $h^+(\mathcal{C}_K) = \max(1, h(\mathcal{C}_K))$ and proves the following

Theorem 1.1. *There exists an effectively computable constant $c > 0$ depending only on $\max(d_1, d_2)$ and $[K : \mathbb{Q}]$ with the property that*

$$\max(|\Delta_1|, |\Delta_2|) \leq c\, h^+(\mathcal{C}_K)^8$$

for all CM-points $(j(\tau_1), j(\tau_2))$ with τ_1, τ_2 imaginary quadratic and discriminant Δ_1, Δ_2 which lie on $\mathcal{C}_K$ but not on any modular curve given by $\Phi_m(X, Y) = 0$ with $m \leq 4\max(d_1, d_2)$.

The version presented above is slightly sharper than Kühne's original version since $8+\epsilon$ is replaced by just 8 so that the appearance of ϵ gets eliminated. One of the main features in Kühne's approach is that it does not rely on a result of Masser on algebraic approximation of Kleins j-invariant. Instead he reduces the problem to linear forms in logarithms, in one case classical logarithms and in a second case elliptic logarithms, an approach which seems to us more natural and aesthetically more pleasing. It emphasizes the moduli aspect of the underlying problem in an elegant way and serves also to work out the general pattern of the problem.

Our approach centers around the Fourier expansion of the Klein j-invariant where the main result is Proposition 2.1 in Section 2.3 which is then used in Section 2.4 to deduce sharp asymptotic inequalities for special points on the given curve. The asymptotic inequalities are turned in Section 2.5 into logarithmic forms, one for the totally degenerate case and here Baker's theory is applied, the other is for the semi-abelian case and here elliptic logarithms appear and the author's theory for group varieties comes into play. In the elliptic case the transformation from points to logarithms is achieved through elliptic logarithms given by the Abel-Jacobi map for the Legendre family.

Unfortunately Kühne's paper contains a nontrivial gap related to estimates for the j-function and a wrong reduction to elliptic logarithmic forms.

These two technical problems can fortunately be fixed with quite some effort (see subsection 2.5 and 2.7 and also the letter of Kühne to the author dated 26/07/2014 included with the authorization of the writer in Appendix B). One of the sources for the gap mentioned above is some misinterpretation of the notion *effectively computable*. This misunderstanding seems to be quite in common and needs some clarification. A constant to be effectively computable means that it is a primitive recursive function depending on a set of initial data. Numerical computations by computer programs are not allowed. The term *effectively computable* does not mean that it is *efficiently computable*. This is a further issue which of course is important but not an aspect which is considered here.

In the second part, in Section 3, we turn to mixed Shimura varieties and establish one very special case of the Pink conjecture, see [35] Conjecture 1.1 and see also section A.4. It deals with finite products of universal elliptic curves over modular curves. We use a result of S. Lang on the variation of heights in families of abelian schemes over a complete regular scheme of dimension 1 over a number field. Lang's result is conditional subject to the existence of what Lang calls *good completion*. It generalizes a result of Tate in [40] on the variation of the Néron-Tate height for families of elliptic curves. We use a fundamental result of Faltings and Chai who obtained a partial result on the existence of such a completion. It turns out that their result is sufficient to deal with the situation which comes up in our approach and which enables us to establish an extension of Tate's theorem on the variation of the Néron-Tate height in families from families of elliptic curves to families of abelian varieties. This is mentioned on p.287 in Tate's paper at the end of the second paragraph as an open problem. Our Theorem 3.2 gives the answer for this question and is effective in principle.

For our application of Theorem 3.2 it becomes crucial to know that the absolute height of special points tends to infinity when the absolute value of the discriminant goes to ∞. This is not obvious and one has to refer to deep results. One way is to use uniform distribution of special points studied by Clozel and Ullmo in [9]. This approach is not known but expected to be effective. Another approach uses a theorem of Colmez [11] on the growth of the Faltings height and this result is effective. However one has to make sure that it remains true if the maximal order of the underlying imaginary quadratic field is replaced by an arbitrary order. Colmez refers to a paper by Nakkajima and Taguchi. Some more work needs to be done to get the necessary estimates for making sure that the Faltings height tends to infinity. The final mosaic stone then is a result of Lang and Néron on the Mordell-Weil group of an abelian variety over a function field of dimension 1.

It has been known since quite some time that the Generalized Riemann Conjecture implies André-Oort, see [22]. It is conceivable that the structure of our proof can be used to deal with the André-Oort conjecture for curves in the Siegel modular variety. This is the reason why we were mainly concerned in developing a structured proof rather than in doing all explicit computations. The possibly only exception is Section 2.3 which deals with the boundary of the moduli space of products of two elliptic curves. Here the calculations are very explicit and in some sense a test case for the corresponding situation in higher dimension, e.g. abelian surfaces as described by Igusa.

It should be also pointed out that the Shimura variety $\mathbb{A}^1 \times \mathbb{A}^1$ can be replaced in both the André-Oort and the Pink case by any product of modular curves. This only requires some additional standard arguments.

These notes are an extended version of a series of lectures which we gave in May and June 2012 when visiting the University of Hong Kong. We thank the Department of Mathematics there for its support and hospitality. The final version has been supported by the FWF-Projekt P26114 of the Austrian Science Fund FWF and partly written up during a visit at the Ernst Schrödinger Institute ESI in Vienna in November and December 2013.

2. Effective André-Oort for $\mathbb{A}^1 \times \mathbb{A}^1$

This first section deals with the situation when the underlying mixed Shimura variety is pure and then a finite product modular curves. There is an obvious extension of Theorem 1.1 to this case, in particular when the modular curves are $\mathbb{A}^1$. Since any product of modular curves is a covering of a product of projective lines it suffices to prove the André-Oort conjecture for $(\mathbb{A}^1)^n$. We confine ourselves with $n = 2$ since the proof carries over very easily to the general case.

2.1. The modular invariant $j(q)$

The Eisenstein series $E_k(\tau)$ for $k \geq 2$ are modular forms of weight $2k$ and can be written as a Fourier series

$$E_k(\tau) = 1 + (-1)^k \frac{4k}{B_k} \sum_{n \geq 1} \sigma_{2k-1}(n) q^n$$

in $q = e^{2\pi i \tau}$ with B_k the k-th Bernoulli number, see Serre [38]. The multiplicative arithmetic function $\sigma_{2k-1}(n) = \sum_{d|n} d^{2k-1}$ can be bounded from above by $2^{\omega(n)} n^{2k-1}$ with $\omega(n) \ll \log(n)/\log\log(n)$ the number of distinct prime factors of n. We also consider the modular Δ-function $\Delta(\tau) = (2\pi)^{12}\eta(\tau)^{24}$ with $\eta(\tau)$ the Dedekind η-function given by

$$\eta(\tau) = e^{\frac{\pi i \tau}{12}} \prod_{n \geq 1} (1 - q^n) = \sum_{n=1}^{\infty} \chi(n) e^{\frac{1}{12}\pi i n^2 \tau}$$

where χ is the Dirichlet character mod 12. We put $F(z) = \eta^{24}$ and get

$$F(z) = \sum_{n \geq 1} \tau(n) q^n;$$

here $\tau(n)$ is the Ramanujan τ-function and with $\tau(n) \ll n^6$. This bound follows from Deligne's work on the Weil conjectures which gives the even better $O(n^{\frac{11}{2}+\epsilon})$. The elliptic modular function $j = \frac{E_2^3}{\eta^{24}}$ can be expressed as a Fourier series

$$\begin{aligned} j(\tau) &= \sum_{n \geq -1} c(n) q^n \\ &= q(\tau)^{-1} + 744 + 196884\, q(\tau) + etc. \end{aligned}$$

with

$$c(n) \leq 6e^{4\pi\sqrt{n}}$$

for $n \geq 1$ as has been shown by Herrmann in [20].

The estimates for the coefficients show that the Fourier expansions of the three functions converge as soon as $\tau \in \mathfrak{H}$. The growth of $j(\tau)$ can be easily estimated. To do so one has just to notice that the coefficients $c(n)$ are all real and non-negative and this implies that $|j(\tau)| \leq j(i\ \mathrm{Im}(\tau))$.

Let $\mathfrak{H}$ be the upper half plane and $\mathfrak{H}^* = \mathfrak{H} \cup \{\infty\} \cup \mathbb{Q}$ the standard compactification. The function $j : \mathfrak{H} \to \mathbb{P}^1$ defined above extends to a holomorphic map from $\mathfrak{H}^*$ to $\mathbb{P}^1$ which maps ∞ to $\infty = (0 : 1) \in \mathbb{P}^1$ and which is modular of weight 0.

2.2. Complex multiplication

The proof of the theorem relies on some basic facts on complex multiplication. In this section we briefly recall the theory as far as it is needed for our purpose. For details we refer to the standard literature, in particular to the monographs of Lang [24], §8.1, Neukirch [31], IV.1 and to Cox [13].

For imaginary quadratic τ we write $D(\tau)$ for the discriminant of the field $k = \mathbb{Q}(\tau)$ and Λ_τ for the lattice $\mathbb{Z} + \mathbb{Z}\tau$. The ring of integers of k can be written as $\mathcal{O}_k = \mathbb{Z} + \mathbb{Z}\,\omega(\tau)$ with $\omega(\tau) = \frac{1}{2}(D(\tau) + \sqrt{D(\tau)})$. The endomorphism algebra of the lattice Λ_τ is an order $\mathcal{O} = \mathbb{Z} + \mathbb{Z}\,f\omega(\tau)$ in $\mathcal{O}_k$ with conductor $\mathfrak{f} = (f)$. A lattice $\Lambda \subset k$ is called a proper $\mathcal{O}$-lattice if its endomorphism algebra is $\mathcal{O}$. A proper $\mathcal{O}$-lattice is the same as a fractional ideal in $\mathcal{O}$. We let $Cl(\mathcal{O})$ be the ideal class group of $\mathcal{O}$, the quotient of the multiplicative monoid of proper $\mathcal{O}$-ideals by the multiplicative monoid of principal $\mathcal{O}$-ideals. It is isomorphic to $Cl_{\mathfrak{f}}(\mathcal{O})$, the quotient of the monoid $I^{\mathfrak{f}}(\mathcal{O})$ of $\mathcal{O}$-ideals prime to $\mathfrak{f}$ by the submonoid $P^{\mathfrak{f}}(\mathcal{O})$ of principal ideals prime to $\mathfrak{f}$ (loc.cit.).

There is a multiplicative bijection between the monoid $I^{\mathfrak{f}}(\mathcal{O})$ and the monoid $I^{\mathfrak{f}}(\mathcal{O}_k)$ of ideals in $\mathcal{O}_k$ prime to $\mathfrak{f}$. It is given by $\mathfrak{a} \mapsto \mathfrak{a}^e := \mathfrak{a}\,\mathcal{O}_k$ and inverse $\mathfrak{a} \mapsto \mathfrak{a}^c := \mathfrak{a} \cap \mathcal{O}$. The image of $P^{\mathfrak{f}}(\mathcal{O})$ in $I^{\mathfrak{f}}(\mathcal{O}_k)$ is the subgroup $P^{\mathfrak{f}}(\mathbb{Z})$ of the group of principal ideals $P^{\mathfrak{f}}(\mathcal{O}_k)$ in $\mathcal{O}_k$ prime to $\mathfrak{f}$ generated by elements α prime to f with $\alpha \equiv a \bmod \mathfrak{f}$ for some $a \in \mathbb{Z}$. The quotient $I^{\mathfrak{f}}(\mathcal{O}_k)/P^{\mathfrak{f}}(\mathbb{Z})$ is called the *ring class group* of $\mathcal{O}$ and is naturally isomorphic to $Cl(\mathcal{O})$.

The ring class group is closely related to the *ray class group modulo* $\mathfrak{f}$. This is the quotient of $I^{\mathfrak{f}}(\mathcal{O}_k)$ by the principal ideals $P^{\mathfrak{f},1}(\mathbb{Z}) \subseteq P^{\mathfrak{f}}(\mathbb{Z})$ generated by elements α prime to $\mathfrak{f}$ with $\alpha \equiv 1 \bmod \mathfrak{f}$. We denote the quotient by $Cl_{\mathfrak{f}}(\mathcal{O}_k)$. One gets an exact sequence

$$0 \to P^{\mathfrak{f}}(\mathcal{O}_k)/P^{\mathfrak{f},1}(\mathbb{Z}) \to Cl_{\mathfrak{f}}(\mathcal{O}_k) \to Cl(\mathcal{O}) \to 0. \tag{2.1}$$

The existence theorem in class field theory associates with the ring class group the ring class field $K_{\mathcal{O}}$ and it associates with the ray class group the ray class field $K_{\mathfrak{f}}$. From (2.1) we deduce via Galois theory that $K_{\mathcal{O}} \subseteq K_{\mathfrak{f}}$ with Galois group $\mathrm{Gal}(K_{\mathfrak{f}}/K_{\mathcal{O}}) = P^{\mathfrak{f}}(\mathcal{O}_k)/P^{\mathfrak{f},1}(\mathbb{Z})$. The ring class field of $\mathcal{O}_k$ is equal to $k(j(\mathcal{O}_k))$ and that of $\mathcal{O}$ is $k(j(\mathcal{O}))$. The ray class field for the modulus $\mathfrak{f}$ is $k(j(\mathcal{O}_k), \tau(1/f; \mathcal{O}_k))$ where τ is the Weber function and $k(j(\mathcal{O}), \tau(w(\tau); \mathcal{O}))$ respectively. It contains the Hilbert class field $k(j(\mathcal{O}_k))$, see [13], Theorem 11.39.

For any ideal $\mathfrak{b}$ in $\mathcal{O}$ prime to $\mathfrak{f}$ we let $(\mathfrak{b}, k)$ be the restriction of the Artin map to the ray-class field $k^{\mathfrak{f}}$. If $\mathfrak{a}_1, \ldots, \mathfrak{a}_{h_{\mathcal{O}}}$ are representatives for the proper ideal classes of $\mathcal{O}$, the values $j(\mathfrak{a}_1), \ldots, j(\mathfrak{a}_{h_{\mathcal{O}}})$ are all algebraic integers and conjugate over k and over $\mathbb{Q}$. The Galois group of the field $k(j(\mathfrak{a}))$ over k for a proper $\mathcal{O}$-lattice $\mathfrak{a}$ is isomorphic to $Cl(\mathcal{O})$ under the map $\mathfrak{b} \mapsto \sigma_{\mathfrak{b}}$ with $\sigma_{\mathfrak{b}}(j(\mathfrak{a})) = j(\mathfrak{b}^{-1}\mathfrak{a})$. Furthermore $\sigma_{\mathfrak{b}}$ is the restriction of $(\mathfrak{b}, k)$ to $k(j(\mathcal{O}))$ so that $j(\mathfrak{a})^{(\mathfrak{b},k)} = j(\mathfrak{b}^{-1}\mathfrak{a})$.

Any proper $\mathcal{O}$-lattice Λ defines an elliptic curve $\mathbb{C}/\Lambda$ with complex multiplication by $\mathcal{O}$ defined over the ring class field $k(j(\mathcal{O}))$ associated to the class group $Cl(\mathcal{O})$ by Galois theory.

2.3. j-Estimates

The j-function has a Fourier expansion $j(\tau) = q^{-1} + \epsilon(q)$ where $q = q(\tau) = e^{2\pi i \tau}$. We use it to introduce the logarithmic distance from a point $x = j(\tau) \in \mathbb{P}^1(\mathbb{C})$ to ∞ as $-\log|j(\tau)|$. The following lemma shows that we can also take $\log|q(\tau)| = -2\pi\ \mathrm{Im}(\tau)$ (see also Corollary 2.1)

Lemma 2.1. *For $\tau \in \mathfrak{H}$ with $\mathrm{Im}(\tau) \geq 1$ one has $|\epsilon(q)| \leq 1728$. There is a formal Fourier expansion*

$$j(\tau)^{-1} - q(\tau) = q^2\,\pi(q)$$

with $\pi(q) = -\epsilon(q)/(1 - q\,\epsilon(q))$. The expansion is convergent if $\mathrm{Im}(\tau) > \frac{1}{2\pi}\log 1728$ and $|\pi\,(q(\tau))\,| \leq 3456$ as soon as $\mathrm{Im}(\tau) > \frac{1}{2\pi}\log 3456$.

Proof. The sum on the right of the identity

$$j(\tau)q(\tau) = 1 + q(\tau)\epsilon(q(\tau))$$

deduced from the Fourier expansion for $j(\tau)$ is formally invertible with inverse

$$(1 + q\,\epsilon(q))^{-1} \;=\; 1 - \frac{q\,\epsilon(q)}{1 + q\,\epsilon(q)}.$$

which gives

$$j(\tau)^{-1} q(\tau)^{-1} \;=\; 1 - \frac{q\,\epsilon(q)}{1 + q\,\epsilon(q)}.$$

For the bound for $\epsilon(q)$ we observe that $|\epsilon(q(\tau))| \leq \epsilon(q(i\,\mathrm{Im}(\tau)))$ and therefore

$$|\epsilon(q)| \leq \sum_{n=0}^{\infty} c(n) e^{-2\pi n\,\mathrm{Im}(\tau)} \leq \sum_{n=0}^{\infty} c(n) e^{-2\pi n} = j(i) - e^{2\pi} \leq 1728.$$

If $\mathrm{Im}(\tau) > \frac{1}{2\pi}\log 1728$ then $1 - |q\,\epsilon(q)| > 0$, moreover $|1 + q\,\epsilon(q)| \geq 1 - |q\,\epsilon(q)| > 1/2$ if $\mathrm{Im}(\tau) > \frac{1}{2\pi}\log 3456$. □

If $\pi\,\mathrm{Im}(\tau) > \log 1728$ the term $q(\tau)\epsilon(q(\tau))$ can be estimated from above by $e^{-\pi\,\mathrm{Im}(\tau)}$ which gives

$$1 - e^{-\pi\,\mathrm{Im}(\tau)} \leq |j(\tau) q(\tau)| \leq 1 + e^{-\pi\,\mathrm{Im}(\tau)}.$$

Taking logarithms leads to

Lemma 2.2. *We have*

$$\log(1 - e^{-\pi\,\mathrm{Im}(\tau)}) \leq \log|j(\tau)| - 2\pi\,\mathrm{Im}(\tau) \leq \log(1 + e^{-\pi\,\mathrm{Im}(\tau)})$$

for $\pi\ \mathrm{Im}(\tau) \geq \log 1728$.

We need also a modification of Lemma 2.1.

Proposition 2.1. *Let $\rho > 0$ be a real number. Then*

$$j(\tau)^{-\rho} = q(\tau)^{\rho} - q(\tau)^{1+\rho}\,\vartheta(q(\tau))$$

with $\vartheta(T)$ a power series in T with $|\vartheta(q(\tau))| \leq 3456\,e^{\rho}$ as soon as $2\pi\ \mathrm{Im}(\tau) \geq \rho + \log 3456$.

Proof. Again we start with the identity $j(\tau)q(\tau) = 1 + q\,\epsilon(q)$ and raise it to the ρ^{th} power

$$j(\tau)^{\rho} q(\tau)^{\rho} = e^{\rho \log(1 + q\,\epsilon(q))}.$$

Since $\log(1+x) = \sum_{k\geq 1}(-1)^{k-1}x^k/k$ it follows that $\exp(\rho\log(1+x)) = 1 + x\pi(x)$ with π a formal power series in x. Since both power series are convergent the series $\pi(x)$ is convergent.

Now again $(1+x\pi(x))^{-1} = 1 - \frac{x\pi(x)}{1+x\pi(x)}$. We deduce that $\frac{1}{\exp(\rho\log(1+x))} = 1 - \frac{x\pi(x)}{1+x\pi(x)}$ which we write as $1 - x\,\psi(x)$ where $\psi(x) = \frac{\pi(x)}{1+x\pi(x)}$ and define $\vartheta(q) = -\epsilon(q)\psi(q\epsilon(q))$ whence

$$j(\tau)^{-\rho} = q(\tau)^{\rho} + q(\tau)^{1+\rho}\,\vartheta(q(\tau)).$$

To work out the estimates we write formally

$$\begin{aligned}\exp(\rho\,\log(1+x)) &= \sum_{n\geq 0}\frac{(\rho\,\log(1+x))^n}{n\,!} = \sum_{n\geq 0}\frac{\rho^n}{n\,!}\Big(\sum_{k\geq 1}\frac{(-1)^{k-1}x^k}{k}\Big)^n\\ &= 1 + x\,\pi(x)\end{aligned}$$

with $\pi(x) = \sum_{n\geq 1}\frac{\rho^n x^{n-1}}{n\,!}\Big(\sum_{k\geq 0}\frac{(-1)^k x^k}{k+1}\Big)^n$ a formal power series in x. Clearly $\sum_{k\geq 0}\frac{|x|^k}{k+1} \leq \frac{1}{1-|x|}$ if $|x|<1$ and therefore

$$|\pi(x)| \leq \rho\sum_{n\geq 1}\frac{(\rho|x|)^{n-1}}{n!}\Big(\frac{1}{1-|x|}\Big)^n \leq \rho\sum_{n\geq 0}\frac{(\rho|x|)^n}{(n+1)!}\Big(\frac{1}{1-|x|}\Big)^{n+1}.$$

The term on the right does not exceed $2\rho\exp(\rho)$ as soon as $|x|\leq 1/2$ and this finally leads to

$$|\pi(x)| \leq \exp(2\rho). \tag{2.2}$$

We choose $\mathrm{Im}(\tau)$ so large that $e^{2\pi\,\mathrm{Im}(\tau)} \geq 3456\,e^{\rho}$ which is satisfied as soon as

$$2\pi\ \mathrm{Im}(\tau) \geq \rho + \log 3456$$

and then $|q||\epsilon(q)| \leq 1/2$. The estimates (2.2) are applied with $x = q(\tau)\epsilon(q(\tau))$ to give

$$\begin{aligned}|\vartheta(q)| &= |\epsilon(q)\psi(q\,\epsilon(q))| = \frac{|\epsilon(q)\pi(q\,\epsilon(q))|}{|1+q\,\epsilon(q)\pi(q\,\epsilon(q))|} \leq 1728\frac{|\pi(q\,\epsilon(q))|}{1-|q\,\epsilon(q)\pi(q\,\epsilon(q))|}|\\ &\leq 3456\,e^{\rho}\end{aligned}$$

as stated in the proposition. □

The subsequent corollary is an easy consequence of the preceding proposition.

Corollary 2.1. *The absolute values of* $j(\tau)^{-1}$ *and* $q(\tau)$ *are related by*

$$\frac{1}{2} \leq \left|\frac{j(\tau)^{-1}}{q(\tau)}\right| \leq 2.$$

as soon as $\mathrm{Im}(\tau) > \frac{1}{2\pi}\log 6912 = 1 \cdot 4070911\cdots$.

Proof. We observe that

$$|q(\tau)| - |j(\tau)^{-1} - q(\tau)| \leq |j(\tau)^{-1}| \leq |q(\tau)| + |j(\tau)^{-1} - q(\tau)|,$$

use that $3456 \leq \frac{1}{2}e^{2\pi\,\mathrm{Im}(\tau)} = \frac{1}{2}|q(\tau)|^{-1}$ and the statement follows after division by $q(\tau)$. ☐

2.4. Puiseux

We embed $\mathbb{A}^1 \times \mathbb{A}^1$ into the product $\mathbb{P}^1 \times \mathbb{P}^1$ of two projective lines over the number field K with $\{\infty\} \times \mathbb{P}^1 \cup \mathbb{P}^1 \times \{\infty\}$ the divisor at infinity and we let $\mathcal{D}_K$ be the divisor at infinity of $\mathcal{C}_K$. Denote by x, y the affine coordinate functions in $\mathbb{A}^2$. At each point o of $\mathcal{D}_K$ there are Puiseux expansions $v = v(u^\rho)$ with ρ a positive rational number and u, v local parameters on $\mathbb{P}^1(\mathbb{C}) \times \mathbb{P}^1(\mathbb{C})$ at o such that $(u, v(u^\rho)) \in \mathcal{C}_K(\mathbb{C})$ and $(u(o), v(o)) = (0,0)$. They parametrize the local analytic branches of $\mathcal{C}_K$ at o. For (u, v) we take

$$(u,v) = \begin{cases} (x - x(o), y^{-1}), & x(o) \neq \infty, y(o) = \infty \\ (x^{-1}, y - y(o)), & x(o) = \infty, y(o) \neq \infty \\ (x^{-1}, y^{-1}), & x(o) = \infty, y(o) = \infty. \end{cases} \tag{2.3}$$

The Puiseux expansions can be written in the form $v = \gamma u^\rho + u^{2\rho}\delta(u^\rho)$ with $\gamma \neq 0$ which after division by u^ρ becomes

$$vu^{-\rho} = \gamma + u^\rho \delta(u^\rho) \tag{2.4}$$

with $\delta(T)$ a convergent power series in the variable T. The coefficient γ is algebraic and has degree at most $[K : \mathbb{Q}]\max(d_1, d_2)$.

Lemma 2.3. *There is an effectively computable constant* $c_1 > 1$ *which depends only on the height and on the degree of* $\mathcal{C}_K$ *such that* $|\delta(u^\rho)| < c_1$ *as soon as* $|u| < c_1^{-1}$.

Proof. The Puiseux expansions are g-functions for which the radius of convergence depends only on the height of the curve. The estimate stated in the Lemma follows then from simple geometric series estimates. □

In particular $\gamma^{-1}u^{-\rho}v = 1 + u^\rho\, O(1)$ and by symmetry we get the same result if we take the Puiseux expansion of u in terms of v. Therefore we may take the expansion which suites us most in the case we are dealing with.

Let $(j(\tau_1), j(\tau_2))$ be a point on $\mathcal{C}_K(\mathbb{C})$ and write $(x(o), y(o)) = (j(\sigma_1), j(\sigma_2))$. Then (2.3) becomes

$$(u(j(\tau_1)), v(j(\tau_2))) = \begin{cases} (j(\tau_1) - j(\sigma_1), j(\tau_2)^{-1}), & j(\sigma_1) \neq \infty, j(\sigma_2) = \infty \\ (j(\tau_1)^{-1}, j(\tau_2) - j(\sigma_2)), & j(\sigma_1) = \infty, j(\sigma_2) \neq \infty \\ (j(\tau_1)^{-1}, j(\tau_2)^{-1}), & j(\sigma_1) = \infty, j(\sigma_2) = \infty. \end{cases}$$

In the case that $j(\sigma_1) = j(\sigma_2) = \infty$ the expansion (2.4) takes the form

$$j(\tau_2)^{-1} j(\tau_1)^\rho = \gamma + j(\tau_1)^{-\rho}\, \delta(j(\tau_1)^{-\rho}).$$

From Lemma 2.3 we deduce that $|\delta(j(\tau_1)^{-\rho})| < c_1$. Together with Corollary 2.1 this shows that for $\mathrm{Im}(\tau_1)$ sufficiently large we have

$$\frac{1}{2}\,|\gamma| \leq |j(\tau_2)^{-1} j(\tau_1)^\rho| \leq 2\,|\gamma|$$

which implies that $|j(\tau_2)|$ and then also $\mathrm{Im}(\tau_2)$ is large. We write $q_i = q(\tau_i)$ for $i = 1, 2$. By virtue of Corollary 2.1 this gives

$$\frac{1}{4}\,|\gamma| \leq |q_2\, q_1^{-\rho}| \leq 4\,|\gamma|. \tag{2.5}$$

In conclusion our discussion can be summarized by the following

Lemma 2.4. *For* $(j(\tau_1), j(\tau_2))$ *on* $\mathcal{C}_K$ *the inequalities*

$$\frac{1}{2^\rho c_1}\,|q_1|^\rho \leq |j(\tau_2)^{-1} j(\tau_1)^\rho - \gamma| \leq 2^\rho c_1\,|q_1|^\rho$$

are satisfied as soon as $\mathrm{Im}(\tau_1) > \frac{1}{2\pi} \log c_1$.

The above statement can be reformulated in terms of the function $q(\tau)$ instead of $j(\tau)$.

Lemma 2.5. *There is an effectively computable constant $c_2 > 0$ such that*

$$|(-1)^{2\tau_2 - 2\rho\tau_1}\gamma^{-1} - 1| < c_2\, e^{-2\pi\rho\,\mathrm{Im}(\tau_1)}.$$

Proof. By Lemma 2.1 and Proposition 2.1 together with Lemma 2.4 and (2.5) we get

$$\begin{aligned} |q_2 - \gamma\, q_1^{\rho}| &\leq |j(\tau_2)^{-1} - \gamma\, j(\tau_1)^{-\rho}| + |j(\tau_2)^{-1} - q_2| + |\gamma|\, |j(\tau_1)^{-\rho} - q_1^{\rho}| \\ &\leq |j(\tau_2)^{-1} - \gamma\, j(\tau_1)^{-\rho}| + 3456|q_2|^2 + |\gamma|\, |q_1|^{1+\rho}. \end{aligned}$$

We divide by $|\gamma\, q_1^{\rho}|$ and note that $q = e^{2\pi i\tau} = e^{2\tau\log(-1)}$ to get

$$|(-1)^{2\tau_2 - 2\rho\tau_1}\gamma^{-1} - 1| = |\gamma^{-1}||q_2 q_1^{-\rho} - \gamma| \leq c_2\, e^{-2\pi\rho\,\mathrm{Im}(\tau_1)}$$

and the statement follows. □

Similarly we deduce in the remaining cases when $j(\sigma_1) \neq \infty$ or $j(\sigma_2) \neq \infty$ that $j(\tau_1) - j(\sigma_1) = \gamma\, j(\tau_2)^{-\rho} + j(\tau_2)^{-2\rho}\delta(j(\tau_2)^{-\rho})$ and $j(\tau_2) - j(\sigma_2) = \gamma\, j(\tau_1)^{-\rho} + j(\tau_1)^{-2\rho}\delta(j(\tau_1)^{-\rho})$ which leads to

Lemma 2.6. *There is an effectively computable constant $c_3 > 0$ such that for $(j(\tau_1), j(\tau_2))$ on $\mathcal{C}$ one has*

$$\frac{1}{2c_3}|q_1|^{\rho} \leq |j(\tau_2) - j(\sigma_2)| \leq 2\, c_3|q_1|^{\rho}$$

or

$$\frac{1}{2c_3}|q_2|^{\rho} \leq |j(\tau_1) - j(\sigma_1)| \leq 2\, c_3|q_2|^{\rho}$$

if $\mathrm{Im}(\tau_1)$ *or* $\mathrm{Im}(\tau_2)$ *exceed* $\frac{1}{2\pi}\log c_1$ *respectively.*

2.5. Elliptic logarithms

The estimates given in Lemma 2.6 can be expressed in terms of the differences $|\tau_i - \sigma_i|$ for $i = 1, 2$. For this we use the Abel-Jacobi map for elliptic curves. The easiest way is to take the family of elliptic curves $\mathcal{E} \xrightarrow{\pi} \mathbb{P}^1$

with level 2 structure given by the Legendre family $y^2 = x(1-x)(\lambda - x)$ with degenerate fibers at 0, 1, ∞. There are morphisms

$$\mathbb{P}^1 \xrightarrow{\tau} \widehat{\Gamma(2)\backslash\mathfrak{H}} \xrightarrow{\kappa} \widehat{\Gamma(1)\backslash\mathfrak{H}} \xrightarrow{j} \mathbb{P}^1$$

with τ and j isomorphisms and κ a ramified covering of degree $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(2)] = 6$ with ramification above $j(e^{2\pi i/3}) = 0$, $j(i) = 1728$ and ∞ with indices 3 at 0 and 2 at 1728 and ∞. The middle terms in the diagram are the standard compactifications of the quotients at the cusps. We deduce that the composition is an ramified covering $\mathbb{P}^1 \to \mathbb{P}^1$ and the cusps of $\widehat{\Gamma(2)\backslash\mathfrak{H}}$ corespond to the points 0, 1 and ∞.

The morphism $\tau : \mathbb{P}^1 \to Y(2) = \widehat{\Gamma(2)\backslash\mathfrak{H}}$, $\lambda \mapsto \tau(\lambda)$, is the period map. It is obtained by choosing the 1-form

$$\omega(\lambda) = \omega(\lambda, x) = \frac{dx}{\sqrt{x(1-x)(\lambda - x)}}$$

on the fiber $\mathcal{E}_\lambda = \pi^{-1}(\lambda)$ together with a basis $\gamma_0(\lambda)$ and $\gamma_1(\lambda)$ for the homology of the fibers $\mathcal{E}_\lambda$ and taking the quotient of the corresponding periods modulo $\Gamma(2)$. The standard way to construct the basis is to take two copies of $\mathbb{P}^1$ and to cut them along the straight lines from 0 to 1 and from λ to ∞ and to glue then crosswise. This gives the two, the upper and the lower, sheets of the Riemann surface.

We choose $1/2 > \delta > 0$ sufficiently small and define for all λ with $\min(|\lambda|, |\lambda - 1|) \geq 3\,\delta$ the path $\gamma_0(\lambda) : [0,1] \to \mathcal{R}(\lambda)$ on the Riemann surface $\mathcal{R}(\lambda)$ beginning on the upper sheet, as

$$\gamma_0(\lambda)(t) = \begin{cases} 2\delta\, e^{8\pi i\, t}, & 0 \leq t \leq \frac{1}{4} \\ 2\delta + (4t-1)(1-4\delta), & \frac{1}{4} \leq t \leq \frac{1}{2} \\ 1 + 2\delta\, e^{(8t-5)\,\pi\, i}, & \frac{1}{2} \leq t \leq \frac{3}{4} \\ 1 - 2\delta - (4t-3)(1-4\delta), & \frac{3}{4} \leq t \leq 1 \end{cases}$$

In a similar way we define the path $\gamma_1 = \gamma_1(\lambda)$ as

$$\gamma_1(\lambda)(t) = \begin{cases} 1 + \delta\, e^{8\pi i\, t}, & 0 \le t \le \frac{1}{4} \\ 1 + \delta + (4t-1)(\lambda - 1 - 2\delta), & \frac{1}{4} \le t \le \frac{1}{2} \\ \lambda + \delta\, e^{\pi\, i\,(8t-5)}, & \frac{1}{2} \le t \le \frac{3}{4} \\ \lambda - \delta - (4t-3)(1 - \lambda + 2\delta), & \frac{3}{4} \le t \le 1 \end{cases}$$

beginning on the lower sheet at $1+\delta$, changing at $t = 1/8$ to the upper sheet and again back at $t = 5/8$ into the lower sheet. The two curves $\gamma_0(\lambda)$ and $\gamma_1(\lambda)$ meet for λ fixed in one single point, the point $1+\delta$ in the upper sheet of the Riemann surface. This shows that they generate the homology $H_1(E_\lambda, \mathbb{Z})$, the path γ_0 being independent of λ whereas the path γ_1 does vary with λ. They have the important property that

$$\inf_{t\in[0,1]} (|\gamma_i(t)|, |\gamma_i(t) - 1|, |\gamma_i(t) - \lambda|) \ge \delta \tag{2.6}$$

for $i = 0,1$ and all λ such that $\min(|\lambda|, |\lambda - 1|) \ge 3\,\delta$ provided that the disc around λ of radius δ does not intersect $\gamma_i(\lambda)$. If the latter happens not to be the case we have to modify γ_i and replace the part of γ_i which lies in the disc by one of the parts of the boundary of the disc cut off by γ_i in such a way that the orientation of the path remains continuous. This ensures that (2.6) holds for the modified paths. The loss coming from the modification depends only linearly on δ.

The form $\omega = \omega(\lambda)$ is holomorphic outside the fibers $\pi^{-1}\{0,1,\infty\}$. We write $\langle \omega, \gamma\rangle = \int_\gamma \omega$ for the period of ω along $\gamma \in H_1(E_\lambda(\mathbb{C}), \mathbb{Z})$. The periods which are associated with our set of data are $W_0(\lambda) = \langle \omega(\lambda), \gamma_0\rangle$ and $W_1(\lambda) = \langle \omega(\lambda), \gamma_1\rangle$ and we define $\tau(\lambda)$ to be $W_1(\lambda)/W_0(\lambda)$ modulo $\Gamma(2)$.

We choose $\epsilon = 3\delta$ and denote by $U(\epsilon)$ the complement in $\mathbb{P}^1(\mathbb{C})$ of the (closed) discs with radius ϵ centered around $\{0,1,\infty\}$. It is contained in the affine line $\mathbb{A}^1(\mathbb{C})$. The following proposition is a key for getting effectivity.

Proposition 2.2. *There exists an effectively computable positive constant c_4 depending only on δ such that for all $\lambda_1, \lambda_2 \in U(\epsilon)$ with $|\lambda_2 - \lambda_1| < \delta$ and with $l^+(\gamma_1(\lambda_1))$ the maximum of* 1 *and the length of $\gamma_1(\lambda_1)$ we have*

$$|W_i(\lambda_2) - W_i(\lambda_1)| < c_4\, l^+(\gamma_1(\lambda_1))\, |\lambda_2 - \lambda_1| \tag{2.7}$$

for $i = 0, 1$.

Proof. For the proof we write $\lambda_2 = \lambda_1 + u$ and estimate the difference

$$W_i(\lambda_1 + u) - W_i(\lambda_1) = \int_\gamma \frac{d}{d\lambda} W_i(\lambda). \tag{2.8}$$

with $\gamma : I = [0,1] \to U(\epsilon)$ a smooth path with $\gamma(0) = \lambda_1$ and $\gamma(1) = \lambda_1 + u$. By a theorem of Ehresmann, see Theorem 4.1.2 in [8], the family $\mathcal{E} \to \mathbb{P}^1$ is locally differentiably trivial over $\mathbb{P}^1 \setminus \{0, 1, \infty\}$ and in particular over γ. As a consequence there exists a canonical horizontal extension γ_i^* of $\gamma_i(\lambda_1)$ over γ such that $\gamma_i^*(\lambda_0) = \gamma_i(\lambda_0)$. The class of the fiber $\gamma_i^*(\lambda)$ for $\lambda \in \gamma$ is homologous to the class of $\gamma_i^*(\lambda_1) = \gamma_i(\lambda_1)$ which is itself homologous to the class of the fiber $\gamma_i(\lambda)$, see [7], p.680 ff. We deduce that[1]

$$\frac{d}{d\lambda} W_i(\lambda) = \int_{\gamma_i(\lambda)} \frac{\partial}{\partial\lambda}\omega(\lambda).$$

It follows that the difference (2.8) can be bounded from above by

$$\sup_{\lambda\in\gamma, t\in\gamma_i(\lambda)} |\frac{\partial}{\partial\lambda}\omega(\lambda,t)| \cdot \sup_{\lambda\in\gamma}(l(\gamma_i(\lambda)) \cdot l(\gamma) \tag{2.9}$$

with $l(\gamma)$ and $l(\gamma_i(\lambda))$ the lengths of γ and $\gamma_i(\lambda)$ respectively. To derive (2.7) we choose γ of minimal length and then $l(\gamma)$ can be bounded from above by a constant multiple of $|\lambda_2 - \lambda_1|$ because we may take as $\gamma(t) = \lambda_1 + tu$ for $0 \le t \le 1$. The statement of the proposition then follows easily.

It remains to deduce an upper bound for $\frac{\partial}{\partial\lambda}\omega(\lambda, t)$ along the paths $\gamma_i(\lambda)$ for $\lambda \in \gamma$ and an upper bound for $l(\gamma_i(\lambda))$ for $\lambda \in \gamma$. This needs to consider

---

[1] *here the calculation:* let $U \subseteq \mathbb{P}^1 \setminus \{0, 1, \infty\}$ be an open disc of small enough radius centered around λ_1 such that $E_U = p^{-1}(U) \overset{\phi}{\longleftarrow} E_{\lambda_1} \times U$ is a trivialization. Then $\gamma_i^*(t) = \phi_*(\gamma_i(\lambda_1), t)$ and $(\xi, \eta) := \phi^*(x, y)$ is independent of t. Also one immediately verifies by going through the explicit formula for $\omega(t)$ that $\frac{d}{dt}\phi^*\omega(t) = \phi^*\frac{d}{dt}\omega(t)$. Furthermore the operators $\frac{d}{dt}$ and $\int_{(\gamma_i^*(\lambda_1), t)}$ commute and therefore our claim results from

$$\begin{aligned} \frac{d}{dt}\int_{\gamma_i(t)} \omega(t) &= \tfrac{d}{dt}\textstyle\int_{\phi_*(\gamma_i(\lambda_1), t)} \omega(t) = \frac{d}{dt}\int_{(\gamma_i(\lambda_1), t)} \phi^*\omega(t) = \\ \int_{(\gamma_i(\lambda_1), t)} \frac{d}{dt}\phi^*\omega(t) &= \textstyle\int_{(\gamma_i(\lambda_1), t)} \phi^* \tfrac{d}{dt}\omega(t) = \int_{\gamma_i(t)} \frac{d}{dt}\omega(t) \end{aligned}$$

where we have used the fact that the integrals depend only on homology classes which are the same for γ_i^* and γ_i.

the derivative with respect to the horizontal parameter λ of the algebraic function $w(\lambda) = \omega(\lambda)/dx$. The length of $\gamma_i(\lambda)$ for $\lambda \in \gamma$ can be easily estimated in terms of the length of $\gamma_i(\lambda_1)$. This is particularly easy for $i = 0$ since then $\gamma_0(\lambda)$ is constant in λ and this shows that $l(\gamma)$ is at most equal to an absolute constant since $\delta < 1/2$ by assumption. For $i = 1$ it is clear that the length of $\gamma_1(\lambda_1)$ and that of $\gamma_1(\lambda)$ for $\lambda \in \gamma$ differ only by $2|u| < 2\delta$ so that $l(\gamma_1(\lambda)) \leq l(\gamma_1(\lambda_1)) + 1 \leq 2\, l^+(\gamma_1(\lambda_1))$.

The upper bound for $\frac{\partial}{\partial\lambda}\omega(\lambda)$ is obtained by first computing the derivative of $\omega(\lambda)$ with respect to λ as

$$\frac{\partial}{\partial\lambda}\omega(\lambda) = -\frac{1}{2}\frac{\omega(\lambda)}{(\lambda - x)}$$

and then estimating the denominator for $\lambda \in \gamma$ both along $\gamma_0(\lambda)$ and along $\gamma_1(\lambda)$ from below by a constant multiple of $\delta^{5/2}$ using (2.6) to get a constant multiple of $\delta^{-5/2}$ as upper bound for $|\frac{\partial}{\partial\lambda}\omega(\lambda)|$. The proposition now follows by putting the three estimates together. $\square$

Our next step is to relate λ_1 and λ_2 to the corresponding j_1 and j_2. Between λ and j there is an algebraic relation of the form

$$j = 2^8 \frac{(1 - \lambda + \lambda^2)^3}{\lambda^2(1-\lambda)^2} \tag{2.10}$$

which can be used to express λ_1 and λ_2 in terms of j_1 and j_2 respectively and then to rewrite the inequality in the proposition in terms of j_1 and j_2 and vice versa. The covering $\mathbb{P}^1 \xrightarrow{q} \mathbb{P}^1$ defined by the algebraic relation (2.10) is ramified only above 0, 1728 and ∞. There the ramification index is 3 at 0 and 2 at the remaining two points. We rewrite the relation (2.10) as an equation $F(j, \lambda) = 0$ for a curve Y of degree 6 and apply again Puiseux. First we express for given $Q = (j_1, \lambda_1) \in Y(\overline{\mathbb{Q}})$ the polynomial F in j and λ as a polynomial $G(j - j_1, \lambda - \lambda_1)$ in $j - j_1$ and $\lambda - \lambda_1$. It has height bounded by a constant multiple of $H(j_1)^6$ where $H(j_1)$ denotes the height of j_1 and with the constant being absolute. A priori the bound would also depend on the height of λ_1 which however, up to an absolute constant, can be estimated using the relation (2.10) from above by the height of j_1. Now we proceed as in the proof of Lemma 2.4. As a result we obtain for the given $Q = (j_1, \lambda_1) \in Y(\overline{\mathbb{Q}})$ an effectively computable disc D around j_1 and a Puiseux expansion $\lambda - \lambda_1 = (j - j_1)^{1/e}\delta((j - j_1)^{1/e})$ for $\lambda - \lambda_1$ in terms of

$(j - j_1)^{1/e}$ with e the ramificationn index at Q. The latter takes the value 1, 2 or 3 as we have seen. The radius r of the disc depends in an effective way only on the height of j_1. It follows that $|\lambda - \lambda_1|$ and $|j - j_1|^{1/e}$ only differ by an effectively computable factor as soon as r is chosen such that $|\delta((j - j_1)^{1/e})| \leq r^{-1/2e}$ and we obtain

$$|\lambda - \lambda_1| \leq r^{-\frac{1}{2e}} |j - j_1|^{\frac{1}{e}}. \tag{2.11}$$

This requires to take $1/r > c_5(H(j_1)^{\varkappa}$ for positive constants c_5 and $\varkappa$ which can be determined effectively. It follows that $|\lambda - \lambda_1| < r^{\frac{1}{2e}}$. We finally observe that the length $l^+(\gamma_i(\lambda_1))$ can also be estimated from above by a constant multiple of $|\lambda_1| + 1$ and that the latter can be replaced by $H(j_1)$ as seen above.

In our application λ_1 is contained in the fiber over j_1 of the covering κ and will by construction be the second coordinate of a point $Q = (\infty, j(\sigma_2))$ on $\mathcal{C}_K$ for some σ_2 in the fundamental domain of $\mathrm{SL}_2(\mathbb{Z})$ with $j(\sigma_2) \neq \infty$. Therefore λ_1 does not belong to the fiber of κ over ∞ which consists of $0, 1, \infty$. The height of $j(\sigma_2)$ can be estimated in terms of the height of the curve $\mathcal{C}_K$ since it is on the intersection of $\mathcal{C}_K$ with $\infty \times \mathbb{A}^1 \subset \infty \times \mathbb{P}^1$. This carries over to λ_1 and implies by Liouville that its distance to any of the points $0, 1, \infty$ can be bounded from below by 4δ as soon as δ^{-1} exceeds a constant multiple of $\exp(c_6 \max(d_1, d_2) h^+(\mathcal{C}_K))$ with c_6 a positive constant which depends only on $[K : \mathbb{Q}]$. We take $r < \delta^{2e}$ and then because of (2.11) all λ with $|\lambda - \lambda_1| < \delta$ are in $U(3\delta)$. In conclusion we get

Corollary 2.2. *There are effectively computable positive constants $c_7 = c_7(r)$ and $\varkappa$ such that for all $j \in D$ there exists λ such that $F(j, \lambda) = 0$ and*

$$|W_i(\lambda) - W_i(\lambda_1)| < c_7 H(j_1)^{\varkappa} |j - j_1|^{1/e} \tag{2.12}$$

for $i = 0, 1$.

2.6. Logarithmic forms

As explained in Section 2.3 and Section 2.4 we have to deal with two cases depending on the behavior of the curve $\mathcal{C}_K$ at infinity. The first case refers to Lemma 2.5 which leads to consider classical linear forms in logarithms. In the second case starting from Lemma 2.6 we arrived in Corollary

2.2 at two linear forms in elliptic logarithms which then lead, as we shall see, to one single linear form of a very special type. What we then need are lower bounds for such linear forms. Since the cases are completely different we have to deal with them separately.

2.6.1. Classical logarithmic forms

In this subsection we briefly recall the result on linear forms in logarithms which we shall use in the final part of the proof of Theorem 1.1. In general, diophantine problems lead to linear forms with rational coefficients. Then one usually uses the bounds given in [5]. We are however very exceptionally in the case where the coefficients are algebraic irrationals lying in an imaginary quadratic field and here the known bounds are a priori not quite as sharp as in the so-called rational case. However there is a special feature which treats the different α_i differently and as a consequence the bound we get is as good as in the rational case as we shall see.

Let $\alpha_1, \ldots, \alpha_n$, all different from 0 and 1, and $\beta_1, \ldots, \beta_n$ be elements in an algebraic number field K,

$$L(z) = L(z_1, \ldots, z_n) = \beta_1 z_1 + \cdots + \beta_n z_n$$

a non-zero linear form with height $h(L)$ and define $h^+(L) = \max(1, h(L))$. We also define $h^+(\alpha) = \max(1, h(\alpha))$ for $\alpha \in K^\times$, write $\Omega = h^+(\alpha_1) \cdots h^+(\alpha_n)$, $\Omega' = \max(e, \Omega/h^+(\alpha_n))$ and further assume that $h^+(\alpha_n) = \max(h^+(\alpha_i))$. Then we have the following

Theorem 2.1. *There exists an effectively computable positive constant c_8 which depends only on n and $[K : \mathbb{Q}]$ such that*

$$\log |L(\log \alpha_1, \ldots, \log \alpha_n)| > -c_8\, h^+(L)\, \Omega\, \log \Omega'. \tag{2.13}$$

unless $L(\log \alpha_1, \ldots, \log \alpha_n) = 0$.

The standard reference for this result is [3]. We shall apply the theorem in the simple case when $n = 2$, $\alpha_1 = -1$, $\alpha_2 = \gamma$, $\beta_1 = \tau_2 - \rho\tau_1$ with τ_1 and τ_2 imaginary quadratic, ρ rational, and $\beta_2 = -1$. In this situation the second order term $\log \Omega'$ disappears because of $h^+(-1) = 1$. Of course this improvement is not really relevant since in this "totally degenerate" case the estimates are much better than in the "semi-degenerate" case to which we come now.

2.6.2. Elliptic logarithmic forms

Let $E = E(g_2, g_3)$ be an elliptic curve over the number field K defined by the equation

$$y^2 = 4x^3 - g_2 x - g_3 \tag{2.14}$$

with g_2 and g_3 in K and $\wp(z) = \wp(z; g_2, g_3)$ the Weierstrass elliptic function associated with g_2 and g_3. We define $h^+(E)$ to be $\max(1, h)$ with h the logarithmic height of the point $(1, g_2^3, g_3^2) \in \mathbb{A}^3(K)$. This is a height of E. Another height is $h^+(j(E)) = \max(1, h(j(E))$. Both heights can be compared and one easily shows that

$$c_9^{-1} h^+(E) \leq h^+(j(E)) \leq c_9\, h^+(E) \tag{2.15}$$

for some effectively computable positive constant c_9. We further choose generators ω_1, ω_2 in the period lattice Λ of E such that $\tau = \omega_2/\omega_1$ is in the fundamental domain of $\mathrm{SL}_2(\mathbb{Z})$. As a consequence the imaginary part $\mathrm{Im}(\tau)$ is bounded from below by $\sqrt{3}/2$ and the absolute value of the real part is bounded from above by $1/2$.

If $\gamma = (\gamma_1, \gamma_2)$ is a point in $(E \times E)(K)$ we let u_i be chosen in the fundamental domain in $\mathbb{C}$ with respect to Λ such that $\gamma_i = (\wp(u_i), \wp'(u_i))$ and write $u = (u_1, u_2)$. The Néron-Tate height $\hat{h}(\gamma)$ of γ is zero if the point is torsion. This will be now assumed. We take a linear form $L(z_1, z_2)$ different from 0 with coefficients in K and introduce the quantities

$$\log \nu_i = \frac{|u_i|^2}{|\omega_i|^2\ \mathrm{Im}(\tau)}$$

and $\log \nu = \max(\log \nu_1, \log \nu_2)$. Also we write $\log^+ x = \max(1, \log x)$ for real $x \geq 0$. Then from Theorem 1.6 in [14] the following

Theorem 2.2. *There exists an effectively computable positive constant c_{10} which depends only on $[K : \mathbb{Q}]$ such that*

$$\begin{aligned}&\log |L(u)| \\ &> -c_{10}(h^+(L) + h^+(E) + \log\log^+ \nu)(h^+(E) + \log\log^+ \nu)^3 \log^+ \nu_1 \log^+ \nu_2.\end{aligned}$$

unless $L(u_1, u_2) = 0$.

can be deduced. In the next section where the proof of Theorem 1.1 is given we shall obtain an upper bound for the right hand side in terms of $\mathcal{C}_K$ and the discriminant Δ of E.

2.6.3. Legendre versus Weierstrass and vice versa

There remains the small problem of comparing Weierstrass elliptic curves $E(g_2, g_3)$ to the Legendre elliptic curves $E(\lambda)$ considered in Section 2.5 with defining equation

$$y^2 = x(x-1)(x-\lambda) \tag{2.16}$$

and to carry over our estimates obtained for the latter to the former (see also [19], ch. IV, §4). Let $K[u, v]$ be the affine coordinate ring for $E(g_2, g_3)$ and $K[x, y]$ that of $E(\lambda)$ and write

$$\begin{aligned} g_2 &= \frac{\sqrt[3]{4}}{3}\left(\lambda^2 - \lambda + 1\right) \\ g_3 &= \frac{1}{27}(\lambda+1)(2\lambda^2 - 5\lambda + 2) \\ \Delta(g_2, g_3) &= \Delta(\lambda) = \lambda^2(\lambda-1)^2 \\ j(g_2, g_3) &= 2^8\frac{(1-\lambda+\lambda^2)^3}{\lambda^2(\lambda-1)^2}. \end{aligned} \tag{2.17}$$

Then the curves $E(g_2, g_3)$ and $E(\lambda)$ are isomorphic with the isomorphism given by

$$\begin{aligned} E(g_2, g_3) &\xrightarrow{\varphi} E(\lambda) \\ \varphi^*(x, y) &= (\sqrt[3]{4}u + \frac{\lambda+1}{3}, v). \end{aligned}$$

The invariant differential forms $\omega(\lambda)$ and $\omega(g_2, g_3) = dx/2y$ are related by

$$\varphi^*\omega(\lambda) = 2\sqrt[3]{4}\,\omega(g_2, g_3).$$

It follows that in Corollary 2.2 the periods $W_i(\lambda_1)$ and $W_i(\lambda_2)$ for the curves $E(\lambda_i)$ for $i = 1, 2$ in the inequality (2.12) can be replaced by the periods $W_i(g_{2,1}, g_{3,1})$ and $W_i(g_{2,2}, g_{3,2})$ associated through (2.17) with the Weierstrass elliptic curves $E(g_{2,i}, g_{3,i})$. This changes the constant c_7 there only by an absolute constant. The same remark also holds for the different heights.

2.7. The proof

The curve $\mathcal{C}_K$ is defined over K and if we replace it by the union $\mathcal{C}$ of its conjugates over $\mathbb{Q}$ we may assume that $K = \mathbb{Q}$. For the proof we assume that $P = (j_1, j_2)$ is a CM-point on $\mathcal{C}$, we denote by Λ_ν the lattice associated with j_ν and write Δ_ν for the discriminant of the lattice Λ_ν. By definition $j_\nu = j(\Lambda_\nu)$ and by the theory of complex multiplication this is an algebraic integer for $\nu = 1, 2$. The lattice Λ_ν is a proper $\mathcal{O}_\nu$-lattice for an order $\mathcal{O}_\nu \subseteq \mathcal{O}_{k_\nu}$ in an imaginary quadratic number field k_ν with discriminant Δ_ν and conductor $\mathfrak{f}_\nu$. The order can be written as $\mathcal{O}_\nu = \mathbb{Z} + \mathbb{Z}\tau_\nu$ with $\tau_\nu = (\Delta_\nu + \sqrt{\Delta_\nu})/2$, see §5.2 in [10], and then $j(\mathcal{O}_\nu) = j(\tau_\nu)$. We denote by $\mathfrak{a}_\nu$ the fractional ideal in $\mathcal{O}_\nu$ associated with Λ_ν and then $j(\mathfrak{a}_\nu) = j_\nu$.

The field $K_\nu = k_\nu(j(\mathcal{O}_\nu)) \supseteq k_\nu$ is the ring class field of the order $\mathcal{O}_\nu$ and as such abelian over k_ν and galois over $\mathbb{Q}$. The same is the case for the compositum $L = K_1K_2$ of K_1 and K_2 with a Galois group $G = Gal(L/\mathbb{Q})$. The restriction of the group G to K_ν is the class group of $\mathcal{O}_\nu$ and acts on K_ν by $j(\mathfrak{a})^{(\mathfrak{b},k_\nu)} = j(\mathfrak{b}^{-1}\mathfrak{a})$ as explained in section 2.2. This shows that the orbit P^G of the CM-point P under G contains a CM-point $Q = (j(\mathcal{O}_1), j(\mathfrak{a}_2))$ which is on $\mathcal{C}$ since the latter is defined over $\mathbb{Q}$. The fractional ideal $\mathfrak{a}_2$ can according to Proposition 5.2.1 in loc. cit. be represented as $a\mathbb{Z}+(b+\sqrt{\Delta_2})\mathbb{Z}$ and then $j(\mathfrak{a}_2) = j(\tau_2)$ with $\tau_2 = \frac{b+\sqrt{\Delta_2}}{2a}$. Further it can be arranged that both of τ_1 and τ_2 are in the fundamental domain and in particular that $\mathrm{Im}(\tau_2) \geq \frac{\sqrt{3}}{2}$. This implies that $-a < b \leq a$ and $|\Delta_2| \geq 12a^2$. It follows that the height $h(\tau_2)$ is at most $c_{11} \log |\Delta_2|$. Without loss of generality we may assume that $|\Delta_2| \leq |\Delta_1|$.

Then we are either in the situation of Lemma 2.5 or Lemma 2.6. In the first case we get a logarithmic form $2(\tau_2 - \rho\tau_1)\log(-1) - \log\gamma$ with γ algebraic. The height of γ can be estimated in terms of the height $h^+(\mathcal{C})$, the height of $\tau_2 - 2\rho\tau_1$ by a constant multiple of $\log|\Delta_1|$ with the constant being absolute and then (2.13) together with Lemma 2.5 leads to

$$\sqrt{|\Delta_1|} \leq C' h^+(\mathcal{C}_K) \log |\Delta_1| \tag{2.18}$$

which is stronger than what we have stated in the theorem unless $\tau_2 = \rho\tau_1$ and $\gamma = 1$. This defines in the universal covering space $\mathfrak{H}\times\mathfrak{H}$ a modular curve and in this case the point $(j(\tau_1), j(\tau_2))$ is on a modular curve as described in the Theorem.

Or the second case applies and then Lemma 2.6 shows that the difference $j(\tau_2) - j(\sigma_2)$ is small. From Corollary 2.2 we deduce that for $i = 0, 1$ the difference $L_i = w_{i,1} - w_{i,2}$ is small for $w_{i,1} = W_i(j(\tau_2))$ and $w_{i,2} = W_i(j(\sigma_2))$. We get elliptic logarithmic forms L_i in $\omega_{i,1}$ and $\omega_{i,2}$ with integer coefficients. Since $j(\tau_2)$ and $j(\sigma_2)$ are algebraic the complex numbers $\omega_{i,1}$ and $\omega_{i,2}$ are periods of elliptic curves defined over a number field. Elimination gives $\omega_{0,2} - \tau_2 \omega_{1,2}$ small and this is a logarithmic form $L(\omega_{0,2}, \omega_{1,2})$ in the periods associated with $j(\sigma_2)$.

In our case the elliptic curves in Section 2.5 coincide with the elliptic curve E with j-invariant $j(\sigma_2)$. We have $u_1 = \omega_{0,2}$ and $u_2 = \omega_{1,2}$ and this leads to $\log \nu_1 = 1/\operatorname{Im}(\sigma_2)$ and $\log \nu_2 = |\sigma_2|/\operatorname{Im}(\sigma_2)$ bounded from above. From (2.15) it follows that the height $h^+(E)$ can be bounded by a constant multiple of $h^+(j(\sigma_2))$ which itself can be bounded in terms of $h^+(\mathcal{C}_K)$, the height of our curve $\mathcal{C}_K$, we finally get from Theorem 2.2 the lower bound

$$|\omega_{0,2} - \tau_2 \omega_{1,2}| > -c_{12}(h^+(\mathcal{C}_K) + h^+(L))\, h^+(\mathcal{C}_K)^3.$$

provided that the linear form is non-zero. From Lemma 2.5 together with Corollary 2.2 we deduce the upper bound

$$|\omega_{0,2} - \tau_2 \omega_{1,2}| < e^{-c_{13}\sqrt{|\Delta_1|}}$$

with c_{13} effective and positive. Since $h^+(L)$ can be bounded up to constant factor by $\log |\Delta_1|$ we find that $|\Delta_1| \leq e$ or

$$\sqrt{|\Delta_1|} < c_{14}(h^+(\mathcal{C}_K) + \log |\Delta_1|) h^+(\mathcal{C}_K)^3$$

for some effectively computable positive constant c_{14} which depends only on $[K : \mathbb{Q}]$. This shows that

$$\min\left(\frac{\sqrt{|\Delta_1|}}{h^+(\mathcal{C}_K)}, \frac{\sqrt{|\Delta_1|}}{\log |\Delta_1|}\right) < c_{15}\, h^+(\mathcal{C})^3.$$

and leads to $|\Delta_1| < c_{15}\, h^+(\mathcal{C}_K)^8$ or $\sqrt{|\Delta_1|}/\log |\Delta_1| < c_{15}\, h^+(\mathcal{C}_K)^3$, and then in both cases to a bound as stated in the theorem. At the end we only have to exclude that the linear form is zero. If so then $\tau_2 = \sigma_2$ and this shows that $j(\tau_2) = j(\sigma_2)$ which means that the point $(j(\tau_1), j(\tau_2))$ lies on the curve $\mathbb{P}^1 \times \{j(\sigma_2)\}$ which is modular but has bi-degree $(1, 0)$ so that in this case $d_1 d_2 = 0$ which was excluded.

3. A Special Case of the Pink Conjecture

In this last section we turn to a particular case of the conjecture of Pink for special points on mixed Shimura varieties (see Appendix A). In order to state our result we let S be a modular curve with function field $k(S)$ defined over an algebraic number field K and $\overline{S}$ a smooth compactification. We write Σ for the boundary $\overline{S} \setminus S$ of S. Let $E \to S$ be the universal elliptic curve over S which always exists after a finite base change. The Néron model $N \to \overline{S}$ of E exists since $\overline{S}$ is a Dedekind scheme and contains $E \to S$ as an open set. The latter is a mixed Shimura variety over the pure Shimura variety S. The boundary Σ of the base consists of finitely many closed points over which N becomes degenerate. We remove the singular loci in the degenerate fibers and take the connected components of the zero sections. It can be arranged by finite base change that the resulting scheme $B \to \overline{S}$ is a semi-abelian subvariety of N with fibers B_s and N_s for $s \in \Sigma$. Their quotients $\Phi_s = N_s/B_s$, the component groups, are finite and étale group scheme over the function field $k(s)$ of s.

The product $B^n \to \overline{S}^{\,n}$ of n copies of B is a semi-abelian variety which contains the mixed Shimura variety $A = E^n \to S^n$. Let T be a smooth projective curve over K and $k(T)$ its function field. We consider a $k(T)$-rational point of A which is not constant. By the Néron property it extends to a morphism $\xi : T \to N^n$ which induces by composition with the projection $N^n \xrightarrow{\pi} \overline{S}^{\,n}$ a morphism $\pi_*\xi = \pi\xi : T \to \overline{S}^{\,n}$. The least common multiple e of the exponents e_s of the groups Φ_s has the property that multiplication $[e]$ by e maps N^n into B^n. We say that ξ is special if $[e]_*\xi$ dominates the closure in B^n of a special point of A and we say that a $\overline{K}$-rational point $\tau : \operatorname{Spec}\overline{K} \to T$ of T is special if $[e]_*\xi\,\tau$ is special on the mixed Shimura variety A. The latter induces a special point $\pi_*\xi\,\tau : \operatorname{Spec}\overline{K} \to S^n$ on the pure Shimura variety S^n which has a Faltings height $h_{Fal}(\pi_*\xi\,\tau)$ since it corresponds to a product of elliptic curves $E_1 \times \ldots \times E_n$ defined over $\overline{K}$ with complex multiplication by orders in imaginary quadratic number fields.

Theorem 3.1. *There exist effectively computable positive constants a and b such that $\xi \in A(k(T))$ is special if and only if there exists either a special point $\tau \in T(\overline{K})$ such that $h_{Fal}(\pi_*\xi\,\tau) > a$ or $\pi_*\xi$ is constant and there exists a special point $\tau \in T(\overline{K})$ such that $ord(\xi\tau) > b$.*

The theorem together with Theorem 3.3.1, the strong Northcott property for elliptic curves, and a theorem of Raynaud shows that the set of special points on T can be effectively determined, see Section 3.1 and Subsection 3.3.1.

In order to relate the theorem to both, Theorem 1.1 and the conjecture of Pink, and to give an example we take for instance the Legendre family (2.16) or the j-family

$$y^2 = 4x^3 - \frac{27\,j}{j-1728}x - \frac{27\,j}{j-1728}\,, \tag{3.1}$$

both considered as elliptic surfaces $\mathcal{E} \xrightarrow{p} \mathbb{P}^1$ in $\mathbb{P}^2 \times \mathbb{P}^1$ with p the restriction of the second projection to $\mathcal{E}$. The singular fibers can be easily determined in the two cases. As above we take a T-rational point ξ on the product $\mathcal{E}^n \to (\mathbb{P}^1)^n$ and a $\overline{K}$-rational point τ on T with $\xi_*\tau$ torsion and assume that the fiber $E_1 \times \ldots \times E_n$ defined by $\xi_*\tau$ has complex multiplication by orders in imaginary quadratic number fields. By definition τ is a special point of T. If the height $h_{Fal}(E_1 \times \ldots \times E_n)$ is sufficiently large then T is special. In our case this means that T either dominates a translate of an elliptic curve in a fiber of $\mathcal{E}^n$ or that T dominates a torsion section over a module curve in $(\mathbb{P}^1)^n$.

Coming back to Theorem 3.1 we observe at once that if ξ is special then the conclusion of the theorem follows. Therefore we only need to show the converse. For this part of the proof of Theorem 3.1 we make some reductions which are needed in order to apply the result in Section 3.3. We replace ξ by $[e]_*\xi$ and $B^n \to \overline{S}^n$ by the fiber products $G = B^n \times_{\overline{S}^n} T \xrightarrow{q} T$ and then $[e]_*\xi : T \to G$ becomes a section. The semi-abelian variety $G \to T$ is an open subscheme of the Néron model $N \to T$ of its generic fiber, see [6], Proposition 3 in §8.1. Further we assume for simplicity that $n = 2$ since the general case can be reduced easily to this special case. The proof is carried out in two steps, the first consists of showing that $[e]_*\xi$ and then also ξ is a torsion section and the second step is an application of Theorem 1.1.

3.1. The special case

In the special case when $\pi_*\xi$ is constant the image of ξ is either a closed point or ξ dominates a curve C in a fiber $E_1 \times E_2 = \sigma^*(\mathcal{E} \times \mathcal{E})$ with $\sigma = \pi_*\xi\tau$. If the image of ξ is a closed point then ξ is constant and this has been excluded. Otherwise the image is a curve in a product $E_1 \times E_2$ of two elliptic curves defined over a number field with each having complex multiplication. Then the genus of C must be at least 1 since an abelian variety does not contain rational curves. If the curve has genus > 1 a theorem of Raynaud, see [36], says that in this case $\xi_*\tau$ has bounded order which we also have excluded by taking b sufficiently large. Therefore $g = 1$ and C is a translate by a torsion point. The result can be made effective, for further details see [21].

3.2. The Faltings-Chai good compactification

As our next step we show that the Néron-Tate height of ξ is zero. Our argument relies on an extension by Lang of a theorem of Silverman on the variation of heights in an elliptic family, see [25], Chapter 12, Corollary 5.4. For applying the result we need the existence of a *good completion* of the Néron model N for A where $A \to U$ is an abelian scheme over an open set $U \subseteq T$ for a smooth projective curve T as before endowed with a polarization L. According to Lang this means that there exists a completion $\overline{N} \to T$ of the Néron model $N \to T$ flat over T which contains the latter as an open subscheme and such that addition $N \times_T N \xrightarrow{m} N$ on the Néron model N extends to an action $N \times_T \overline{N} \xrightarrow{\overline{m}} \overline{N}$, together with a relatively ample invertible sheaf $\overline{L}$ on $\overline{N}$ extending a polarization L of A. The existence of such a good completion seems to be still an open problem. However a partial answer has been given by Faltings and Chai and follows from Theorem 1.13 on the compactification of the universal abelian variety $\mathcal{A} \to \overline{A_g}$ over the compactification $\overline{A_g}$ of the moduli stack A_g of principally polarized abelian varieties, see [17], Ch.VI.

Let $G \xrightarrow{q} T$ be the open subgroup scheme of N from above, a semi-abelian scheme over T which by convention means that its generic fiber is abelian and that all its degenerate fibers are semi-abelian varieties. The existence of a completion $\overline{G} \stackrel{\overline{q}}{\to} T$ of G proper over T and smooth over $\mathbb{Z}$ now follows from Theorem 1.13 mentioned above. The semi-abelian scheme $G \to T$ operates on $\overline{G}$ and it is contained in $\overline{G}$ as a dense open subscheme. There exists a relatively ample invertible sheaf $\overline{L}$ on $\overline{G}$ which restricts to $\mathcal{O}(2)$, the canonical symmetric ample invertible sheaf on A which gives twice the principal polarization.

The difference to a good completion of the Néron model $N \to T$ is that our compactification is only a good compactification of $G \to T$ and does not necessarily contain the Néron model. However we know that the least common multiple e of the exponents of the component groups of the fibers is finite and effectively computable. It follows that for each section $\xi : T \to N$ the section $[e]_*\xi$ factors through $G \hookrightarrow \overline{G}$. Since at the end we are only interested in the Néron-Tate height of a section we may replace ξ by $[e]_*\xi$. For our purposes this property is as good as to have a good completion as we shall see. As a consequence we may assume that the section ξ itself factors.

The construction depends on a set of data. One fixes a free abelian group X of rank g and chooses a smooth $GL(X)$-admissible polyhedral cone decomposition $\{\sigma\}$ of the cone $C = C(X) \subseteq B(X)$ of positive semi-definite symmetric bilinear forms in the space of all bilinear forms on X with rational radical. One also considers the space $\widetilde{B}(X) = B(X) \times X$ and in $\widetilde{B}(X)_{\mathbb{R}} = \widetilde{B}(X) \otimes \mathbb{R}$ the cone $\widetilde{C} = \widetilde{C}(X)$ of elements (b, l) such that l vanishes on the radical of b. The group $\widetilde{GL}(X) = GL(X) \ltimes X$ acts on $\widetilde{C}$. Then one chooses a smooth $\widetilde{GL}(X)$-admissible polyhedral cone decomposition $\{\tau\}$ of $\widetilde{C}$ over the cone decomposition $\{\sigma\}$ of C such that each $\sigma \times \{0\}$ is a τ. The construction of the invertible sheaf $\overline{L}$ depends on the choice of a principal polarization-function $\widetilde{\phi}^{\,2}$ on $\widetilde{C}$. All these data can in principal be given effectively.

3.3. Variation of heights

As already mentioned the proof of Theorem 3.1 makes use of an extension by Lang of techniques introduced by Tate in his article on the variation

[2]for details see [17], Ch.VI, §1, in particular Definition 1.5

of the absolute height in a family of elliptic curves $\mathcal{E} \xrightarrow{\pi} T$ over a smooth absolutely irreducible projective curve, see [40]. In our case we have constructed a good completion $\overline{G} \xrightarrow{\overline{q}} T$ of $G \xrightarrow{q} T$ together with an ample invertible sheaf $\overline{L}$. Furthermore we have assumed that the section ξ factors through G. We obtain absolute height functions $h_\eta = h_{\overline{L}_\eta}$ and $h_t = h_{\overline{L}_t}$ associated with this datum on the generic fiber $\overline{G}_\eta$ and on the fibers $\overline{G}_t$ respectively for every point $\operatorname{Spec} \overline{K} \xrightarrow{t} T$. They are defined with respect to the restrictions $\overline{L}_\eta$ and $\overline{L}_t$ of the invertible sheaf $\overline{L}$ to the fibers. Correspondingly we obtain Néron-Tate heights $\widehat{h}_\eta$ and $\widehat{h}_t$ with associated quadratic forms q_η and q_t defined on the non-degenerate fibers. We choose an invertible sheaf M of degree 1 on T and deduce[3]

Theorem 3.2. *The quadratic forms q_t and q_η are related by*

$$q_t(\xi(t)) = q_\eta(\xi(\eta))\, h_M(t) + O(h_M(t)^{1/2}) + O(1). \tag{3.2}$$

Since $\xi(\tau)$ is torsion it follows that $q_t(\xi(\tau))$ is 0 and this implies that $q_\eta(\xi(\eta)) = 0$ provided that $h_M(\tau)$ can be taken sufficiently large. In the next subsections we shall show in two different ways that this can indeed be assumed. Then we need only to show that $\xi(\eta)$ is a torsion point as soon as its Néron-Tate height $h_\eta(\xi(\eta))$ is zero.

3.3.1. The strong Northcott property

Our heights are absolute heights and for applying Theorem 3.2 we need that the set of special points has the strong Northcott property. This means that the subset of special points of bounded absolute height is finite.

Theorem 3.3 (Strong Northcott Property)**.** *There are only finitely many special points on $\mathfrak{H}/\Gamma$ with bounded absolute height.*

It remains to establish the strong Northcott property of the set of special points on $\mathbb{P}^1$. There are two ways to see that. One is to use the paper of Clozel and Ullmo [9] on the equidistribution of Hecke points or one can also apply a very interesting result of Colmez[4] about the Faltings height of CM elliptic curves.

[3] see Corollary 5.4 in Chapter 12 in [25]

[4] see [11], Théorème 1.

Theorem 3.4 (Colmez [11]). *There are effectively computable constants a and $b > 0$ such that for square-free positive integers d the elliptic curve E_d with complex multiplication by the maximal order in $\mathbb{Q}(\sqrt{-d})$ is bounded below by $h_{Fal}(E_d) \geq a + b \log d$.*

However this is not quite what we need since it gives the desired result only in the case when the ring of endomorphisms is the maximal order in the CM-field. If not one has to modify and build in the conductor. In the next two subsections we shall give two different proofs, one based on Clozel-Ullmo and the other refers to Colmez.

3.3.2. Clozel-Ullmo and Duke equi-distribution

Let K be an imaginary quadratic field with field discriminant d_K and ring of integers $\mathcal{O}_K$. The group of classes of invertible ideals in $\mathcal{O}_K$ is finite and has order $h(\mathcal{O}_K)$, the class number of $\mathcal{O}_K$. We denote by $\mathcal{O}_{K,f}$ the order in $\mathcal{O}_K$ with conductor f and discriminant $d = f^2 d_K$. We also write $X(1)$ for $\mathrm{SL}_2(\mathbb{Z})\backslash\mathfrak{H}$ and $\mathcal{C}_0\left(X(1)\right)$ for the space of piecewise continuous functions which tend to 0 at infinity and we introduce the Poincaré measure $d\mu_0 = \frac{3}{\pi}\frac{dx\,dy}{y^2}$. The normalization is chosen such that $\int_{X(1)} d\mu_0 = 1$. Since by elementary integration $\int_{X(1)} \frac{dx\,dy}{y^2} = \frac{\pi}{3}$ we get the normalization factor as chosen.

In general the corresponding normalization factor can be read off from the formula[5]

$$\frac{1}{2\pi}\int_{\Gamma\backslash\mathfrak{H}^*} \frac{dx\,dy}{y^2} = 2g - 2 + m + \sum_1^r \left(1 - \frac{1}{e_i}\right)$$

for a compact Riemann surface $\Gamma\backslash\mathfrak{H}^*$ with Γ a Fuchsian groups of the fist kind, with e_i the orders of the inequivalent elliptic points of Γ and m the number of inequivalent cusps. In our case $g = 0, m = 1, r = 2, e_1 = 2, e_2 = 3$. Then the right hand side becomes $1/6$ and this leads to the normalization factor $3/\pi$. We also write $\Lambda_{K,f}$ for the set of points with complex multiplication by $\mathcal{O}_{K,f}$.

[5]see Theorem 2.20 in [39].

In their equi-distribution paper, see [9], Theorem 2.4, Clozel and Ullmo extended a difficult result of Duke [15] from the case of principal order $\mathcal{O}_K$ to that of an arbitrary order $\mathcal{O}_{K,f}$ with conductor f to obtain

Proposition 3.3. *For any function $\phi \in \mathcal{C}_0(X(1))$ we have*

$$\frac{1}{h(\mathcal{O}_{K,f})} \sum_{y \in \Lambda_{K,f}} \phi(y) \to \int_{X(1)} \phi(x) d\mu_0(x)$$

if the discriminant $d = d(\mathcal{O}_{K,f})$ of the order $\mathcal{O}_{K,f}$ tends to ∞.

Proof. See [9], Theorem 2.4. □

As a consequence of the proposition we find that for given $\epsilon > 0$ there exists a positive constant $d(\epsilon, \phi)$ such that

$$-\epsilon + \int_{X(1)} \phi(x) d\mu_0(x) \leq \frac{1}{h(\mathcal{O}_{K,f})} \sum_{y \in \Lambda_{K,f}} \phi(y) \leq \int_{X(1)} \phi(x) d\mu_0(x) + \epsilon \quad (3.3)$$

as soon as $d \geq d(\epsilon)$. From Section 2.3 we deduce that $\frac{\log |j(\tau)|}{2\pi \operatorname{Im}(\tau)} \to 1$ for $t = \operatorname{Im}(\tau) \to \infty$. It follows that for given $\epsilon > 0$ there exists $r(\epsilon) > 0$ such that $\big| \log |j(\tau)| - 2\pi \operatorname{Im}(\tau) \big| < \epsilon$ as soon as $t > r(\epsilon)$. We choose $\epsilon > 0$ and apply the Proposition to the function $\phi = \chi(r) \cdot \log^+ |j(\tau)|$ where $\chi(r)$ for given $r > 0$ is the characteristic function of the set of $\tau = s + it \in \mathcal{F}$ such that $r(\epsilon) \leq t \leq r$. On observing the well-known fact (see Section 2.2) that $j(\mathcal{O}_{K,f})$ is an algebraic integer, and therefore there are no non-archimedean contributions to the height, one gets

$$\begin{aligned} h(j(\mathcal{O}_{K,f})) &= \frac{1}{h(\mathcal{O}_{K,f})} \sum_{y \in \Lambda_{K,f}} \log^+ |j(y)| \\ &\geq \frac{1}{h(\mathcal{O}_{K,f})} \sum_{y \in \Lambda_{K,f}} \phi(y). \end{aligned}$$

Using the j-estimates given in Section 2.3 we deduce that

$$\begin{aligned} -\epsilon + 2\pi \int_{X(1)} \chi(r) \operatorname{Im}(x) d\mu_0(x) &\leq \int_{X(1)} \chi(r) \log |j(x)| d\mu_0(x) \\ &\leq 2\pi \int_{X(1)} \chi(r) \operatorname{Im}(x) d\mu_0(x) + \epsilon \end{aligned}$$

as soon as $\mathrm{Im}(\tau)$ is sufficiently large. We calculate $\int_{X(1)} \chi(r)\,\mathrm{Im}(x)d\mu_0(x)$ as

$$\frac{3}{\pi}\int_{-1/2\leq s\leq 1/2}\int_{r(\epsilon)\leq t\leq r} ds\,dt/t = \frac{3}{\pi}\log r - \frac{3}{\pi}\log r(\epsilon)$$

and obtain

$$h(j(\mathcal{O}_{K,f})) \geq 6(1-\epsilon)\log r$$

for fixed r as above as soon as $d(\mathcal{O}_{K,f}) \to -\infty$. It follows that for given $R>0$ there are only finitely many orders $\mathcal{O}_{K,f}$ for K imaginary quadratic such that $h(j(\mathcal{O}_{K,f})) < R$ and this implies the strong Northcott property.

3.3.3. Colmez' height bound

As mentioned already above we cannot directly apply the lower bound given by Colmez for the Faltings height in terms of the discriminant of the elliptic curve since the curve needs not to have the maximal order of the imaginary quadratic field as its ring of endomorphisms. Instead we have to deal with the case of an arbitrary order $\mathcal{O}_{K,f}$ of conductor f. This is done by comparing the heights of two isogeneous elliptic curves and here we may use the well-known comparison result for abelian varieties established by Faltings in his celebrated finiteness paper[6]. To state the result let L be a number field and for i=1,2 let $A_i \xrightarrow{p_i} \mathrm{Spec}(\mathcal{O}_L)$ be semi-abelian schemes with proper generic fiber. We also denote by $\mathrm{Spec}(\mathcal{O}_L) \xrightarrow{\epsilon} A_i$ the zero section, by $A_1 \xrightarrow{\phi} A_2$ an isogeny and write $G = \ker(\phi)$. The isogeny induces an injection

$$\phi^* : \omega_{A_2/\mathcal{O}_L} \to \omega_{A_1/\mathcal{O}_L}$$

and one sees that

$$|\omega_{A_2/\mathcal{O}_L}/\phi^*\omega_{A_1/\mathcal{O}_L}| = |\epsilon^*\Omega^1_{G/\mathcal{O}_L}|.$$

An easy calculation shows that

$$h_{Fal}(A_2) = h_{Fal}(A_1) + \frac{1}{2}\log(\deg(\phi)) - \frac{1}{[L:\mathbb{Q}]}\log(|\epsilon^*\Omega^1_{G/\mathcal{O}_L}|). \qquad (3.4)$$

[6]see §5 in [16], in particular Lemma 5

This is applied to the elliptic curves $E_d = \mathbb{C}/f\mathcal{O}_K$ and $E_{d,f} = \mathbb{C}/\mathcal{O}_{K,f}$ and to the natural isogeny $E_{d,f} \to E_d$ which has degree f. In our case $\epsilon^*\Omega^1_{G/\mathcal{O}_L} = \mathcal{O}_L/\mathfrak{d}_G$ with $\mathfrak{d}_G$ the absolute different of G over L. One gets

$$|\mathcal{O}_L/\mathfrak{d}_G| = \prod_{p|f} p^{e(p)\frac{[L:\mathbb{Q}]}{2}}$$

with

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

where $n = \mathrm{ord}_p(f)$ [7]. Clearly $e(p) \leq \frac{4}{p}$ which gives $e(p)\log p \leq 4/e$. We denote by $\omega(n)$ the number of distinct prime factors of n. By analytic number theory $\omega(n) = O(\log n/\log\log n)$ [8] and therefore

$$\sum_{p|f} e(p)\log p \leq (4/e)\,\omega(f) = O(\log f/\log\log f).$$

This together with the lower bound for $h_{Fal}(E_d)$ by Colmez given in Theorem 3.4 inserted in (3.4) gives the existence of effectively computable constants a and $b > 0$ such that

$$h_{Fal}(E_{d,f}) \geq a + b\,\log|d(\mathcal{O}_{K,f})|\,.$$

The strong Northcott property now follows by comparing the j-height and the Faltings height[9].

3.4. Proof of Theorem 3.1

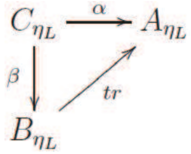
From Section 3.3 we know that $\xi(\eta)$ is a rational point in A_η, the generic fiber of the abelian variety A underlying G, with $\widehat{h}_\eta(\xi(\eta)) = 0$. Now we are in a situation as has been studied by Lang and Néron, see Theorem 1 in [27]. They determined the Mordell-Weil group of an abelian variety defined over a function field in terms of the group of rational points of the abelian variety over $K(T)$ and its trace over K. The $(K(T), K)$-trace (B, tr) of A is an

[7]see §0 in [30] for the notation and for an explanation that $e(p)\frac{[L:\mathbb{Q}]}{2}$ is an integer and §1 for the calculation of $|\mathcal{O}_L/\mathfrak{d}_L|$.

[8]see §22.10 and §22.11 in [18].

[9]see [12], Ch. X, Prop. 2.1.

abelian variety B defined over K together with a homomorphism $B_\eta \xrightarrow{tr} A_\eta$. It has the property that if $L \supseteq K$ is an extension of K which is free from $K(T)$, C an abelian variety defined over L and $C_{\eta_L} \xrightarrow{\alpha} A_{\eta_L}$ a homomorphism then there is a homomorphism $C \xrightarrow{\beta} B_L$ such that



is commutative; here $\eta_L = \eta \otimes_K L$ and subscript means base change. The main result of Lang and Néron in the function field case says that the group $A(K(T))/trB(K)$ is finitely generated. One further knows, see [25] Theorem 1.6, Ch. III, §1, that the property $\widehat{h}_\eta(\xi(\eta)) = 0$ is equivalent to $\xi(\eta) \in A(K(T))_{tor} + trB(K)$. Since $\xi(\tau) \in A_\tau(\overline{K})_{tor}$ by assumption we conclude that $\xi(\eta) \in A(K(T))_{tor} + trB(K)_{tor}$ and this proves the theorem.

Remarks. Our proof of Theorem 3.1 shows that the constant c is effectively computable. However it depends on the very deep result on *good completions* by Faltings and Chai. Their result is constructive and could be made effective in principal. To work this out in a conceptual way would be a very valuable and highly non-trivial contribution to the theory. Also the Weil height of the curve $\xi(T)$ can be determined as well as its order as a point in the Mordell-Weil group $E^n(K(S))$.

Appendix A. Pink's conjecture

In this section we shall briefly state and explain the conjecture of Pink which generalizes the André-Oort conjecture from Shimura varieties to mixed Shimura varieties. The main source for this section are the articles [34] and [35] of Pink. For further details we refer to these papers or to [28].

1. Mixed Shimura varieties

Let $\mathbb{S} := \mathcal{R}_{\mathbb{C}/\mathbb{R}}\mathbb{G}_m$ be the real torus obtained by Weil restriction from $\mathbb{C}$ to $\mathbb{R}$ introduced by Deligne. It has the property that there is a natural isomorphism $\mathbb{S}(\mathbb{R}) \simeq \mathbb{C}^\times$. The Deligne torus, as it is called, can be used to define Hodge structures through representation theory. Namely it is the same to give a representation $h : \mathbb{S} \to \mathrm{GL}(V_{\mathbb{R}})$ in a real vector space $V_{\mathbb{R}}$ or to give a $\mathbb{R}$-Hodge structure on $V_{\mathbb{R}}$.

Let P be a connected linear algebraic group over $\mathbb{Q}$ and consider the set $\mathrm{Hom}(\mathbb{S}_{\mathbb{C}}, \mathrm{P}_{\mathbb{C}})$ on which $P(\mathbb{C})$ acts on the left by composition with the inner automorphism $int(p) : p' \mapsto pp'p^{-1}$ for $p \in P(\mathbb{C})$. By X^+ we denote an orbit under the action of the subgroup $P(\mathbb{R}) \cdot U(\mathbb{C})$ with U a subgroup of the unipotent radical W of P which is normal in P. The pair (P, X^+) is called a connected *mixed Shimura datum* if it satisfies a number of additional conditions, partly related to mixed Hodge structures as introduced by Deligne. If in addition P is reductive then (P, X^+) is a connected pure Shimura datum.

Sufficiently small congruence subgroups $\Gamma \subseteq P(\mathbb{Q})$ act freely on X^+ and then the quotient $\Gamma\backslash X^+$ becomes a complex manifold which has a natural structure of a quasi-projective algebraic variety. The variety is called a connected *mixed Shimura variety* associated to the datum (P, X^+) and to Γ. It is called a *pure Shimura* variety if P is reductive. Morphisms between Shimura varieties, mixed or pure, are defined in an obvious way by going through the Shimura data.

Let V be a finite dimensional representation of P over $\mathbb{Q}$ such that the associated rational mixed Hodge structure has type $\{(-1,0),(0,-1)\}$, in other words it is that of an abelian variety. We choose a Γ-invariant lattice Γ_V in $V_{\mathbb{R}}$. Then the semi-direct product $P \ltimes V_{alg}$ of the vector group $V_{alg} = V \otimes_{\mathbb{Q}} \mathbb{G}_a^{\dim V}$ with P together with the conjugacy class $X^+ \ltimes \Gamma_V \subseteq \mathrm{Hom}(\mathbb{S}_{\mathbb{C}}, \mathrm{P}_{\mathbb{C}} \ltimes \mathrm{V}_{\mathrm{alg},\mathbb{C}})$ generated by X^+ defines a new Shimura datum $(P \ltimes V_{alg}, X^+ \ltimes V_{\mathbb{R}})$. The projection $\pi : P \ltimes V_{alg} \to P$ then induces a Shimura epimorphism

$$A := \Gamma \ltimes \Gamma_V \backslash X^+ \ltimes \Gamma_V \xrightarrow{[\pi]} \Gamma\backslash X^+ =: S.$$

The mixed Shimura variety A obtained by this construction is a family of abelian varieties over S. In general a mixed Shimura variety is a torus bundle over a polarized abelian scheme over a Shimura variety (see [28], Chap. VI, Theorem 1.6).

2. Generalized Hecke operators

Let $S = \Gamma\backslash X^+$ be a connected mixed Shimura variety and $\varphi \in \mathrm{Aut}(P)$. Such a set of data leads to the diagram

$$S = \Gamma\backslash X^+ \xleftarrow{[id]} \Gamma \cap \varphi^{-1}(\Gamma)\backslash X^+ \xrightarrow{[\varphi]} \Gamma\backslash X^+ = S$$

which is called generalized Hecke operator T_φ associated with φ. By definition a Hecke operator is a correspondence on S and we define for any set $\Sigma \subseteq S$ the translate of Σ as

$$T_\varphi(\Sigma) = [\varphi]([\mathrm{id}]^{-1}(\Sigma)).$$

The image of any morphism $S' \xrightarrow{\varphi} S$ of Shimura varieties or of any generalized Hecke operator is called a *special subvariety*. If its dimension is zero then we call it *special point*.

3. Our special case

In our case $P = \mathrm{GL}_2(\mathbb{R})^+$ and $X^+ = \mathfrak{H}$ and we get the Shimura variety $X(\Gamma) = \Gamma\backslash\mathfrak{H}$. If in addition we take the tautological representation $V = \mathbb{Q}^2$ in the above construction we get as a mixed Shimura variety the universal elliptic curve $E \xrightarrow{[\pi]} \Gamma\backslash\mathfrak{H}$ (for further details, see also [29]). The special subvarieties of E are E itself when the dimension is 2, torsion sections and fibers of $[\pi]$ in dimension 1 and torsion points in fibers over CM-points in S. The special subvarieties of $E \times E$ are easily determined.

4. The conjecture (see [35])

Let S be a mixed Shimura variety over the field of complex numbers $\mathbb{C}$. By definition an irreducible component of a mixed Shimura subvariety of S, or of its image under a Hecke operator, is called a special subvariety of S. Consider any irreducible closed subvariety $Z \subseteq S$. Since any intersection of special subvarieties is a finite union of special subvarieties, there exists a unique smallest special subvariety containing Z which is called the special closure of Z and denoted by S_Z. We call the dimension of S_Z the amplitude of Z, and the codimension of Z in S_Z the defect of Z. The defect measures how far Z is away from being special; in particular Z is special if and only if its defect is zero. Moreover Z is called Hodge generic if S_Z is an irreducible component of S, that is, if Z is not contained in any special subvariety of codimension > 0. For any point $s \in S$ the amplitude and the defect of $\{s\}$ coincide and are called the amplitude of s. The points of amplitude zero in S are precisely the special points in S. Moreover s is called Hodge generic if $\{s\}$ is Hodge generic.

Conjecture 1. Consider a mixed Shimura variety S over $\mathbb{C}$, an integer d, and a subset Σ of points of amplitude d. Then any irreducible component Z of the Zariski closure of Σ has defect d.

References

1. Y. André, *G-functions and Geometrie*, Aspekte der Mathematik, Vol. **E13**, Vieweg, 1989.

2. Y. André, Finitude des couples d'invariants modulaires singuliers sur une courbe algébrique plan non modulaire, *J. reine angew. Math.*, **505** (1998), 203-208.

3. A. Baker, The theory of linear forms in logarithms, in Transcendence theory: advances and applications (Proc. Conf., Univ. Cambridge, Cambridge, 1976), Academic Press, London (1977), 1-27.

4. A. Baker and G. Wüstholz, *Logarithmic Forms and Diophantine Geometry*, Cambridge University Press, 2007.

5. A. Baker and G. Wüstholz, Logarithmic forms and group varieties, *J. reine angew. Math.*, **442** (1993), 19-62.

6. S. Bosch, W. Lütkebohmert and M. Raynaud, *Néron Models*, Ergebnisse der Mathematik und ihrer Grenzgebiete, 3. Folge, Band 21, Springer, 1990.

7. E. Brieskorn and H. Knörrrer, *Plane Algebraic Curves*, Birkhäuser, 2012.

8. J. Carlson, S. Müller-Stach and C. Peters, *Period Mappings and Period Domains*, Cambridge University Press, 2003.

9. L. Clozel and E. Ullmo, *Equidistribution des Points de Hecke*, Contributions to automorphic forms, geometry, and number theory, John Hopkins Univ. Press (2004), Baltimore, MD, 193-254.

10. H. Cohen, *A course in computational algebraic number theory*, Graduate Texts in Mathematics **138**, Springer-Verlag, 1993.

11. P. Colmez, Sur la hauteur de Faltings des variétés abéliennes à multiplication complexe, *Compositio Math.*, **111** (1998), 359-369.

12. G. Cornell and J. H. Silverman (editors), *Arithmetic Geometry*, Springer, 1986.

13. D. A. Cox, *Primes of the Form* $x^2 + ny^2$, John Wiley & Sons, 1989.

14. S. David and N. Hirata-Kohno, Linear forms in elliptic logarithms, *J. Reine Angew. Math.*, **628** (2009), 37-89.

15. W. Duke, Hyperbolic Distribution Problems and Half-Integral Weight Maass Forms, *Invent. Math.* **92** (1988), 73-90.

16. G. Faltings, Endlichkeitssätze für Abelsche Varietäten, *Invent. Math.* bf 73, 349-366.

17. G. Faltings, C.-L. Chai, *Degeneration of Abelian Varieties*, Ergebnisse der Mathematik und ihrer Grenzgebiete, 3. Folge, Band 22, Springer Verlag, 1990.

18. G. H. Hardy, E. M. Wright, *An Introduction to the Theory of Numbers*, Clarendon Press, 1979.

19. R. Hartshorne, *Algebraic Geometry*, Graduate Text in Mathematics, Springer Verlag, 1997.

20. O. Herrmann, Über die Berechnung der Fourierkoeffizinten der Funktion $j(\tau)$, *Journ. reine u. angew. Math.*, **274/275** (1975), 187-195.

21. M. Hindry, Autour d'une conjecture de Serge Lang, *Invent. Math.*, **94** (1988), 575-603.

22. B. Klingler and A. Yafaev, The André-Oort conjecture, *Ann. Math.*, **180** (2014), 867-952.

23. L. Kühne, An effective result of André-Oort type I, *Ann. Math.*, **179** (2012), 652-671.

24. S. Lang, *Elliptic Functions*, Graduate Texts in Mathematics **112**, Springer-Verlag, 1987.

25. S. Lang, *Fundamentals of Diophantine Geometry*, Springer-Verlag, 1983.

26. S. Lang [Ed.], *Number Theory III*, Encyclopedia of Mathematical Sciences, **60**, Springer-Verlag, 1991.

27. S. Lang and A. Néron, Rational Points of Abelian Varieties Over Function Fields, *Amer. J. Math.*, **81**, No. 1 (1959), 95-118.

28. J. S. Milne, Canonical Models of (Mixed) Shimura Varieties and Automorphic Vector Bundles, *Automorphic Forms, Shimura Varieties, and L-Functions*, Academic Press (1990), 248-241.

29. B. Moonen, An introduction to Mumford-Tate groups, `http://www.math.uzh.ch/fileadmin/conferencesHomepages/shimuraVarieties_05_2004/Lecture_Notes_Moonen.pdf`

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

31. J. Neukirch, Algebraic Number Theory, *Grundlehren* **322** (1999), Springer-Verlag.

32. R. Noot, Correspondances de Hecke, action de Galois et la conjecture de André-Oort, *Séminaire Bourbaki, année* 2004-2005, *exposé* 942 *Astérisque*, **307** (2006), 165-197.

33. F. Oort, Canonical liftings and dense sets of CM points, Arithmetic Geometry (Ed. Fabrizio Catanese), *Symp. Math.*, **37** (1997), Cambridge University Press, 228-234.

34. R. Pink, A combination of the conjectures of Mordell-Lang and André-Oort, Geometric methods in algebra and number theory, Progress in Mathematics **253**, Birkhäuser (2005), 251-282.

35. R. Pink, A common generalization of the conjectures of André-Oort, Manin-Mumford, and Mordell-Lang, Preprint (2005), `http://www.math.ethz.ch/~pinkri/publications.html`

36. M. Raynaud, Courbes sur une variété et points de torsion, *Invent. math.* **71** (1983), 207-233.

37. Th. Scanlon, A proof of the André-Oort conjecture via mathematical logic, Séminaire Bourbaki 63ème année, 2010-2011, n^o 1037.

38. J.-P. Serre, A course in Arithmetic, *Graduate Texts in Mathematics* **7**, Springer Verlag, 1973.

39. G. Shimura, *Introduction to the Arithmetic Theory of Automorphic Functions*, Princeton University Press, 1994.

40. J. Tate, Variation of the Canonical Height of a Point Depending on a Parameter, *Amer. J. Math.*, **105** (1983), No. 1, 287-294.

41. E. Ullmo, A. Yafaev, Galois orbits and equidistribution of special subvarieties: towards the André-Oort conjecture, *Ann. Math.*, **180** (2014), 823-865.

42. G. Wüstholz, Algebraische Punkte auf analytischen Untergruppen algebraischer Gruppen, *Ann. Math.*, **129** (1989), 501-517.

Appendix B. Kühne's letter to the author

Weil am Rhein, 28/07/2014 (revised: 25/08/2014)

Dear Gisbert,

I (once again) had a closer look at the two problems you mentioned in our recent discussions (and which we have already discussed back in 2012) about my Annals article [9].

I. There is indeed a mistake in the way I invoked linear forms in elliptic logarithms: To be precise, I cite the result of David and Hirata-Kohno [5] on linear forms in elliptic logarithms in a wrong – too naive – way. This error concerns Section 2.4, which demands correction, and of course the argument on p. 665 needs a corresponding adjustment (details below) as well.

Before dwelling on any details, let me mention that this error concerns neither the statements of my article nor does it present a serious obstruction for mathematicians in this field. In fact, the respective part of my article is not as original as are the other parts and can already be found in André's article [1]. In contrast to my article, André uses Masser's effective transcendence measure for $j(\tau)$ [12], which – at least formally – circumvents linear forms in elliptic logarithms. For aesthic reasons, i.e. in order to make the multiplicative case and the elliptic one more look alike, I decided to use the result of [5] instead, and slightly failed.

My misunderstanding (see my presentation of their result as Proposition 1 on p. 657 in [9]) is that you must not take

$$\exp_{\mathcal{E}_\tau}(z) = \big(1 : \wp_\tau(z) : \wp'_\tau(z)\big),$$

where $\wp_\tau$ is the Weierstrass $\wp$-function with period lattice $\mathbb{Z} + \mathbb{Z}\tau$, but

$$\exp^{\mathrm{int}}_{\mathcal{E}_\tau}(z) = \big(1 : \wp_{g_2,g_3}(z) : \wp'_{g_2,g_3}(z)\big), \text{ (int is for ‘intelligent’)},$$

where $\wp_{g_2,g_3}$ is the Weierstrass $\wp$-function for the usual $g_2, g_3 \in \overline{\mathbb{Q}}$ you read off from a given equation

$$Y^2 = 4X^3 - g_2X - g_3$$

of $\mathcal{E}_\tau$ (see p. 40 in [5] for these conventions). As everybody – except me at that time – knows there is a considerable difference; for the period lattice is now $\mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$ and τ does merely appear as quotient ω_2/ω_1.

Indeed, it is strictly necessary to work with both elliptic logarithms ω_1 and ω_2 of $0_{\mathcal{E}_\tau}$ because τ only appears as their quotient ω_2/ω_1. What I wrote in my article, however, would mean that you could prove Schneider's famous result on the j-invariant just by showing the transcendence of a single elliptic logarithm, as τ is a logarithm of $0_{\mathcal{E}_\tau}$ for the ‘normalization’ employed in my article [9]. In this way, a ‘one logarithm result’ (as given by Siegel in [15]) would satisfy, which is not the case...

I hope this describes the problem quite precisely so that I may now come to its solution: One may proceed as André did, using Masser's result in the very explicit version of Faisant and Philibert [6, Théorème 2]; this should also improve the $8+\varepsilon$ in my main Theorem 2 to $6+\varepsilon$. For checking, it suffices to note that this exponent is just

$$2 \cdot (\text{exponent of } \log(H) \text{ in the lower bound for } \log|\Lambda| \text{ on p. 665 of [9]}) + \varepsilon.$$

However, my favorite solution is this one: First, copy Proposition 1 in [9] for $\mathcal{E}_\tau \times \mathcal{E}_\tau$ instead of $\mathbb{G}_a \times \mathcal{E}_\tau$ from the very same article [5] (omit the $\mathbb{G}_a$-part there). Second, you evaluate the linear form $\mathcal{L}$ on p. 665 not at

$$\underline{u} = (1, \tau_0) \text{ with } \exp_{\mathcal{E}_\tau}(\tau_0) = 0,$$

but at

$$\underline{u} = (\omega_1, \omega_2) \text{ with } \exp^{\mathrm{int}}_{\mathcal{E}_\tau}(\omega_1) = \exp^{\mathrm{int}}_{\mathcal{E}_\tau}(\omega_2) = 0 \text{ and } \omega_2/\omega_1 = \tau_0.$$

This corrects the argument straightforwardly. Quantitatively, this modification worsens the exponent $8+\varepsilon$ in Theorem 2 to $12+\varepsilon$ but I consider this a neglectable problem since you may actually do better with the result of [6]. In addition, there might be room for further improvement by balancing the various parameters in [5] in a more intelligent way than I perceive right now.

II. As I recall you objected also vividly to my claim of effectivity for the absolute constant c_3 (resp. c_4) introduced on the sixth line from below on p. 663 (resp. the fifth line from below on p. 664) in [9]. The affected assertion of *loc. cit.* boils down to the following claim:

Claim A. There is an effectively computable constant $C_1 > 0$ such that

$$|j(\tau) - 1728| \geq C_1 |\tau - i|^2 \text{ and } |j(\tau)| \geq C_1 |\tau - \zeta_6|^3 .$$

(resp. **Claim B.** There is an effective constant $C_2 > 0$ such that

$$|j'(\tau)|^{-1} \leq C_2 \min\left\{|j(\tau)|^{-2/3}, |j(\tau) - 1728|^{-1/2}\right\}$$

for all $\tau \in \overline{\mathcal{F}}$ – the closure of the standard fundamental domain of $\mathrm{SL}_2(\mathbb{Z})$.)

There is certainly no quarrel about whether such constants exist by the arguments in *loc. cit.* so that at most their effectivity is disputable. By the time of writing [9], I was very convinced that absolute constants related to evaluations of the j-invariant are effectively computable and did not give much thought to these questions at all. In fact, it still seems to me that the j-invariant is well-posed for machine computations and that one might use these ultimatively.

Nevertheless, I do not want to follow this line of thought here and instead communicate to you another approach, parts of which I obtained as recently as last week. First, it is clear that complex analysis, i.e. the excessive use of contour integration, can be used to obtain Claim A from

(1) an explicit upper bound for $|j(\tau)|$ in certain regions $\{C_3^{-1} < \mathrm{Im}(\tau) < C_3\}$ for sufficiently large C_3, and
(2) an explicit lower bound for $|j''(i)|$ and $|j'''(\zeta_6)|$, $\zeta_6 = \exp(2\pi i/3)$.

I guess we agree that (1) is not an issue since explicit bounds can be found at various places in the literature (see e.g. [3, Section 2]).

As (2) is concerned, there is a way to determine $j''(i)$ and $j'''(\zeta_6)$ explicitly and hence their absolute values with sufficient precision. An explicit expression for $j'''(\zeta_6)$ can be found on p. 150 of [4], whence the idea... By definition (cf. [16, Section I.7]),

$$j(\tau) \;=\; 2^6 3^3 \frac{E_4(\tau)^3}{E_4(\tau)^3 - E_6(\tau)^2},$$

where

$$E_{2k}(\tau) \;=\; 1 - \frac{4k}{B_{2k}} \sum_{n=1}^{\infty} \sigma_{2k-1}(n) q^n,\; k \geq 1.$$

(At this point, one is obliged to remark that E_2 is not a modular form due to the lack of absolute convergence in its defining series.) Here, as usual $q = e^{2\pi i \tau}$, $\sigma_k(n) = \sum_{d|n} d^k$ and B_{2k} is the $2k$-th Bernoulli number. Set

$$\theta = \frac{1}{2\pi i}\frac{d}{d\tau} = q\frac{d}{dq}.$$

There are classical (Ramanujan's PQR-) relations (see [10, Theorem X.5.(iii)]):

$$\theta E_2 = E_2^2/12 - E_4/12,$$
$$\theta E_4 = E_2 E_4/3 - E_6/3,$$
$$\theta E_6 = E_2 E_6/2 - E_4^2/2.$$

Using these relations, a (lengthy) computation shows that

$$\theta j = -2^6 3^3 \frac{E_4^2 E_6}{E_4^3 - E_6^2},$$

$$\theta^2 j = 2^5 3^2 \frac{3E_4^4 + 4E_4 E_6^2 - E_2 E_4^2 E_6}{E_4^3 - E_6^2},$$

and

$$\theta^3 j = 2^3 3 \frac{18 E_2 E_4^4 + 24 E_2 E_4 E_6^2 - 3E_2^2 E_4^2 E_6 - 95 E_4^3 E_6 - 16 E_6^3}{E_4^3 - E_6^2}.$$

Contrary to my expectation, these three expressions are rather simple since the denominator has a very simple derivative, namely

$$\theta(E_4^3 - E_6^2) = 3E_4^2 \theta E_4 - 2E_6 \theta E_6 = E_2(E_4^3 - E_6^2),$$

causing a massive cancellation.

Consequently, (2) reduces to an evaluation of the Eisenstein series E_2, E_4, and E_6 at i (resp. ζ_6). I would like to give this in more detail than strictly necessary since I do not know a reliable source in the literature. Consider the non-zero holomorphic differential $\frac{dX}{Y}$ on the complex elliptic curve given by

$$Y^2 = 4X^3 - g_2(\tau)X - g_3(\tau),$$

where $g_2(\tau) = \frac{(2\pi)^4}{2^2 3} E_4(\tau)$ and $g_3(\tau) = \frac{(2\pi)^6}{2^3 3^3} E_6(\tau)$ (see again [16, Section I.7]). Its period lattice is $\mathbb{Z} + \mathbb{Z}\tau$. For any complex $\omega_1 \neq 0$, the map

$$\varphi_{\omega_1} : (x, y) \longmapsto (\omega_1^2 x, \omega_1^3 y)$$

defines an isomorphism with the complex elliptic curve

$$Y'^2 = 4X'^3 - g_2' X' - g_3'$$

with $g_2' = \omega_1^{-4} g_2(\tau)$ and $g_3' = \omega_1^{-6} g_3(\tau)$ (cf. [11, p. 17]). Furthermore,

$$\varphi_{\omega_1}^* \left(\frac{dX'}{Y'} \right) = \frac{\omega_1^{-2} dX}{\omega_1^{-3} Y} = \omega_1 \frac{dX}{Y}$$

and hence the periods of $\frac{dX'}{Y'}$ form a lattice $\mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$, $\omega_2 = \omega_1 \tau$. In this way, we obtain

$$E_4(\tau) = \frac{2^2 3}{(2\pi)^4} g_2(\tau) = \frac{2^2 3}{(2\pi)^4} \omega_1^4 g_2' \text{ and } E_6(\tau) = \frac{2^3 3^3}{(2\pi)^6} \omega_1^6 g_3'.$$

As it is not needed, I omit the evaluation of E_2. Though it is not a modular form, a similar formula can be derived by using the non-holomorphic differential $\frac{XdX}{Y}$ instead of $\frac{dX}{Y}$ (cf. [8, Section I.2] for the necessary background).

Now, (2) reduces to a computation of periods for the complex elliptic curves with $j = 0$ and $j = 1728$. To evaluate E_4 and E_6 at $\tau = i$, we consider

$$Y^2 = 4X^3 - 4X,$$

i.e. $g_2' = 4$, $g_3' = 0$. Furthermore, by substituting $t = u^{-1/2}$ and using [2, Theorems 2.1.2, 2.2.3 and Corollary 2.2.6] we obtain

$$\begin{aligned} \omega_1 = 2\int_1^\infty \frac{dt}{(4t^3 - 4t)^{1/2}} &= \int_1^\infty \frac{dt}{t^{1/2}(t^2-1)^{1/2}} = \\ &= \frac{1}{2}\int_0^1 u^{-3/4}(1-u)^{-1/2} du = \frac{1}{2} B(1/4, 1/2) = \\ &= \frac{\Gamma(\frac{1}{2})\Gamma(\frac{1}{4})}{2\Gamma(\frac{3}{4})} = \frac{\pi^{1/2} \cdot \Gamma(\frac{1}{4})}{2\sqrt{2}\pi\Gamma(\frac{1}{4})^{-1}} = \frac{\Gamma(\frac{1}{4})^2}{\sqrt{8\pi}} \end{aligned}$$

and thus

$$E_4(i) = \frac{2^2 3}{(2\pi)^4} \omega_1^4 g_2' = \frac{2^4 3}{(2\pi)^4} \frac{\Gamma\left(\frac{1}{4}\right)^8}{64\pi^2} = \frac{3}{2^6} \frac{\Gamma\left(\frac{1}{4}\right)^8}{\pi^6} \text{ and } E_6(i) = 0.$$

Similarly, for evaluation at $\tau = \zeta_6$, we consider

$$Y^2 = 4X^3 - 4,$$

i.e. $g_2 = 4$, $g_3 = 0$. In addition, by substituting $t = u^{-1/3}$, using Legendre's Duplication Formula [2, (2.3.1)] for $z = 1/6$ and Euler's Reflection Formula for $z = 1/3$ [2, Theorem 2.2.3] we obtain this time

$$\begin{aligned} \omega_1 &= 2\int_1^\infty \frac{dt}{(4t^3 - 4)^{1/2}} = \int_1^\infty \frac{dt}{(t^3-1)^{1/2}} \\ &= \frac{1}{3}\int_0^1 u^{-5/6}(1-u)^{-1/2} du = \frac{1}{3} B(1/6, 1/2) = \frac{\Gamma(\frac{1}{6})\Gamma(\frac{1}{2})}{3\Gamma(\frac{2}{3})} = \frac{\Gamma(\frac{1}{3})^3}{2^{4/3}\pi} \end{aligned}$$

(the $2^{8/3}$ in the corresponding formula in [13] is apparently a typo) and therefore

$$E_4(\zeta_6) = 0 \text{ and } E_6(\zeta_6) = \frac{2^3 3^3}{(2\pi)^6}\omega_1^6 g_3 = \frac{3^3}{2^9}\frac{\Gamma(\frac{1}{3})^{18}}{\pi^{12}}.$$

In conclusion,

$$j''(i) = (2\pi i)^2 \theta^2 j(i) = 2^5 3^3 (2\pi i)^2 E_4(i) = -2\cdot 3^4 \frac{\Gamma\left(\frac{1}{4}\right)^8}{\pi^4}$$

and

$$j'''(\zeta_6) = (2\pi i)^3 \theta^3 j(\zeta_6) = 2^7 3 (2\pi i)^3 E_6(\zeta_6) = -2\cdot 3^4 i \frac{\Gamma(\frac{1}{3})^{18}}{\pi^9}.$$

Finally, let me come to Claim B. Here, we are solely interested in points $\tau \in \overline{\mathcal{F}}$ such that $|j'(\tau)|$ is arbitrary small. Since $|j'(\tau)| \to \infty$ if $\mathrm{Im}(\tau) \to \infty$, this means that we may freely assume

$$\overline{\mathcal{F}} \cap \{\mathrm{Im}(\tau) < C_4\}$$

for some (effective) $C_4 > 0$. For such a region, the function

$$E_4(\tau)^3 - E_6(\tau)^2 = \frac{2^6 3^3}{(2\pi)^{12}}\Delta(\tau) = \frac{2^6 3^3}{(2\pi)^{12}} q \prod\nolimits_{n=1}^{\infty} (1-q^n)^{24}$$

(see [16, Section I.7]) is bounded from above by C_5 and from below by C_5^{-1} for some effectively computable constant $C_5 > 1$; for the above infinite product is absolutely convergent. Now,

$$(2\pi)^{12}\left|\frac{E_4^2(\tau)E_6(\tau)}{\Delta(\tau)}\right| = |\theta j(\tau)| = \frac{1}{2\pi}\left|j'(\tau)\right|$$

and thus

$$|E_4^2(\tau)||E_6(\tau)| < C_6 \cdot \left|j'(\tau)\right|, \; C_6 = \frac{C_5}{(2\pi)^{13}}.$$

Since $\left|E_4(\tau)^3 - E_6(\tau)^2\right| \geq C_5^{-1}$ we have

$$\text{(i) } |E_4(\tau)| \geq \left(\frac{1}{2C_5}\right)^{1/3} \text{ or (ii) } |E_6(\tau)| \geq \left(\frac{1}{2C_5}\right)^{1/2}.$$

In case (i),

$$|E_6(\tau)| < (2C_5)^{2/3} C_6 \cdot \left|j'(\tau)\right|$$

and hence

$$\frac{1}{2^{4/3}C_5^{7/3}C_6^2}\cdot|j(\tau)-1728|\le\frac{1}{2^{4/3}C_5^{7/3}C_6^2}\cdot\left|\frac{E_6(\tau)^2}{E_4(\tau)^3-E_6(\tau)^2}\right|\le\left|j'(\tau)\right|^2.$$

In case (ii),

$$|E_4(\tau)|\ <\ 2^{1/4}C_5^{1/4}C_6^{1/2}\cdot\left|j'(\tau)\right|^{1/2},$$

and

$$|j(\tau)|\ \le\ 1728\frac{|E_4(\tau)|^3}{|E_4(\tau)^3-E_6(\tau)^2|}<3456C_5^{7/4}C_6^{3/2}\cdot\left|j'(\tau)\right|^{3/2}.$$

We conclude

$$\frac{1}{231C_5^{7/6}C_6}\cdot|j(\tau)|^{2/3}<\left|j'(\tau)\right|$$

and again our reasoning prevails.

II$^{\text{bis}}$. Let me also recall that after you mentioned your doubts for the first time in August 2012, I actually did another sort of brute force computations and I think I did succeed in this tedious search for an explicit c_3 – without acquiring explicit knowledge on $j''(i)$ or $j'''(\zeta_6)$. However, I did not see the simple argument for Claim B given above. In fact, this work was abandoned once it became clear to us that hypergeometric functions provide also a solution that fits in quite well with the proof given in [9]. In my sketchy notes from November 2012, I used the hypergeometric function ${}_2F_1(\frac{1}{2},\frac{1}{2},1;z)$. This means I worked implicitly with the modular curve $Y(2)$ (and not with $Y(1)$). In fact, this is probably the best documentated case in literature (see [7, Section 9.6]) and provides very nice expressions.

As far as I understood, your recent preprint is a similar solution but avoids these explicit expressions in an elegant way. It seems to me that it also resolves all controversies about the effectivity of both constants c_3 and c_4 from [9].

III. [...]

Best wishes,

Lars

References

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

2. Richard Beals and Roderick Wong, *Special functions*, *Cambridge Studies in Advanced Mathematics* 126, Cambridge University Press, 2010.

3. Yuri Bilu and Pierre Parent, Runge's method and modular curves, *Int. Math. Res. Not.*, **9** (2011), 1997-2027.

4. Yuri Bilu, David Masser, and Umberto Zannier, An effective "theorem of André" for CM-points on a plane curve, *Math. Proc. Cambridge Philos. Soc.*, **154** (2013), 145-152.

5. Sinnou David and Noriko Hirata-Kohno, Linear forms in elliptic logarithms, *J. Reine Angew. Math.*, **628** (2009), 37-89.

6. Alain Faisant and Georges Philibert, Quelques résultats de transcendance liés à l'invariant modulaire j, *J. Number Theory* **25** (1987), 184-200.

7. Dale Husemöller, *Elliptic curves* (second edition), *Graduate Texts in Mathematics* 111, Springer-Verlag, 2004.

8. Nicholas Katz, p-adic Interpolation of real analytic Eisenstein series, *Ann. of Math.* (2), **104** (1976), 459-571.

9. Lars Kühne, An effective result of André-Oort type, *Ann. of Math.* (2), **176** (2012), 651-671.

10. Serge Lang, *Introduction to Modular Forms*, *Grundlehren der mathematischen Wissenschaften* 222, Springer-Verlag, 1976.

11. Serge Lang, *Elliptic functions*, *Graduate Texts in Mathematics* 112, Springer-Verlag, 1987.

12. David Masser, *Elliptic functions and transcendence*, *Lecture Notes in Mathematics* 437, Springer-Verlag, 1975.

13. Yuri Nesterenko, Modular functions and transcendence questions, *Sbornik Mathematics*, **187** (1996), 1319-1348.

14. Theodor Schneider, Transzendenzuntersuchungen periodischer Funktionen II, *J. Reine Angew. Math.*, **172** (1934), 70-74.

15. Carl Ludwig Siegel, Über die Perioden elliptischer Funktionen, *J. Reine Angew. Math.*, **167** (1932), 62-69.

16. Joseph H. Silverman, *Advanced Topics in the Arithmetic of Elliptic Curves* (corrected second printing), *Graduate Texts in Mathematics* 151, Springer-Verlag, 1999.