J. Eur. Math. Soc. 17, 2379–2416 © European Mathematical Society 2015

DOI 10.4171/JEMS/560

D. Masser · U. Zannier

Torsion points on families of simple abelian surfaces and Pell's equation over polynomial rings

(with an appendix by E. V. Flynn)

Received December 8, 2013

Abstract. In recent papers we proved a special case of a variant of Pink's Conjecture for a variety inside a semiabelian scheme, namely for any curve inside anything isogenous to a product of two elliptic schemes. Here we go beyond the elliptic situation by settling the crucial case of any simple abelian surface scheme defined over the field of algebraic numbers, thus confirming an earlier conjecture of Shou-Wu Zhang. This is of particular relevance in the topic, also in view of very recent counterexamples by Bertrand. Furthermore there are applications to the study of Pell equations over polynomial rings; for example we deduce that there are at most finitely many complex t for which there exist $A, B \neq 0$ in $\mathbb{C}[X]$ with $A^2 - DB^2 = 1$ for $D = X^6 + X + t$. We also consider equations $A^2 - DB^2 = c'X + c$, where the situation is quite different.

Keywords. Torsion point, abelian surface scheme, Pell equation, Jacobian variety, Chabauty's theorem

1. Introduction

Motivated by recent work on unlikely intersections, we consider here the following conjecture to be found in our recent article [MZ2].

Conjecture. *Let $\mathcal{S}$ be a semiabelian scheme over a variety defined over $\mathbb{C}$, and denote by $\mathcal{S}^{[c]}$ the union of its semiabelian subschemes of codimension at least c. Let $\mathcal{V}$ be an irreducible closed subvariety of $\mathcal{S}$. Then $\mathcal{V} \cap \mathcal{S}^{[1+\dim \mathcal{V}]}$ is contained in a finite union of semiabelian subschemes of $\mathcal{S}$ of positive codimension.*

This is a variant of a conjecture stated by Pink [Pin] in 2005, which generalised the Zilber Conjectures [Zi] of 2002 to schemes.

In [MZ2] (see also [MZ1] for a short version) we verified this conjecture in a special case where $\mathcal{S}$ is the fibred square of the standard Legendre elliptic family, with coordinates

D. Masser: Mathematisches Institut, Universität Basel, Rheinsprung 21, 4051 Basel, Switzerland; e-mail: David.Masser@unibas.ch

U. Zannier: Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy; e-mail: u.zannier@sns.it

E. V. Flynn: Mathematical Institute, University of Oxford, 24–29 St. Giles, Oxford OX1 3LB, United Kingdom; e-mail: flynn@maths.ox.ac.uk

Mathematics Subject Classification (2010): 11G10, 14K15, 14K20, 11G50, 11G30, 14H40

(X_1, Y_1), (X_2, Y_2), and $\mathcal{V}$ is the curve defined by $X_1 = 2$, $X_2 = 3$. This amounted to the finiteness of the set of complex numbers $\lambda \neq 0, 1$ such that the points

$$\big(2, \sqrt{2(2-\lambda)}\big), \quad \big(3, \sqrt{6(3-\lambda)}\big) \tag{1.1}$$

both have finite order on the elliptic curve E_λ defined by $Y^2 = X(X-1)(X-\lambda)$.

In [MZ3] we generalised the result to any x-coordinates defined over an algebraic closure of $\mathbb{C}(\lambda)$; of course then the y-coordinates are also defined over this closure. (See the paper [BD] of Baker and DeMarco for an analogue in the context of algebraic dynamics.) It turns out that this is equivalent to the Conjecture above with $\mathcal{S}$ isogenous to the product of two isogenous elliptic schemes and $\mathcal{V}$ a curve.

In [MZ4] we further generalised these results to any product of two elliptic schemes, whether isogenous or not.

Here we settle the case of any simple abelian surface scheme defined over the field $\overline{\mathbb{Q}}$ of all algebraic numbers. Together with the previous results this will easily imply the following result.

Theorem. *Let $\mathcal{A}$ be an abelian surface scheme over a variety defined over $\overline{\mathbb{Q}}$, and let $\mathcal{V}$ be an irreducible closed curve in $\mathcal{A}$. Then $\mathcal{V} \cap \mathcal{A}^{[2]}$ is contained in a finite union of abelian subschemes of $\mathcal{A}$ of positive codimension.*

This also confirms a conjecture stated in 1998 by Zhang [Zh, Remark 4a, p. 224]. Recently Bertrand [Bert3] discovered a surprising counterexample when the surface scheme is an extension of an elliptic scheme by the multiplicative group $\mathbb{G}_\mathrm{m}$, which is not abelian. Thus it is reassuring to know that no such surprises exist for the abelian case. In a work [BMPZ] with him and Pillay we have also shown that his are essentially the only counterexamples for semiabelian surfaces. So this work completes the analysis of the above Conjecture for schemes of relative dimension 2 over $\overline{\mathbb{Q}}$. See also the second author's book [Za, pp. 77–80]. And Harry Schmidt has investigated extensions of an elliptic scheme by the additive group $\mathbb{G}_\mathrm{a}$ (which are not even semiabelian). In this connection see also the work [CMZ] with Corvaja.

In [MZ3] and [MZ4] we could treat schemes defined over $\mathbb{C}$ not just $\overline{\mathbb{Q}}$, so that becomes a natural problem here too; there are several very promising approaches involving specialization to the above Theorem.

From [MZ4] we can assume that $\mathcal{A}$ is not isogenous to the product of two elliptic schemes. We will soon see that the base variety can be assumed to be irreducible of dimension at most one. If it is a point, then $\mathcal{A}$ is constant and we retrieve the classical result of Manin–Mumford type in the special situation under consideration. In fact we will appeal to the classical result to eliminate this case.

As in our previous papers we can give simple examples of our theorem for base curves. Thus we get the finiteness of the set of complex numbers

$$\lambda \neq 0, 1, -1, i, -i, \frac{-1+\sqrt{-3}}{2}, \frac{-1-\sqrt{-3}}{2} \tag{1.2}$$

such that the pair of points

$$\big(2, \sqrt{2(2-\lambda)(2-\lambda^2)(2-\lambda^4)}\big), \quad \big(3, \sqrt{6(3-\lambda)(3-\lambda^2)(3-\lambda^4)}\big) \tag{1.3}$$

on the curve defined by

$$Y^2 = X(X-1)(X-\lambda)(X-\lambda^2)(X-\lambda^4) \tag{1.4}$$

give—via the unique point at infinity on (1.4)—a point of finite order on the Jacobian (compare with (1.1)).

We could have similar examples with a polynomial of degree 6 on the right of (1.4), as the genus remains 2. But we could then replace (1.3) by the two points at infinity, where the matter has been well-known since Abel [Ab] (see also Chebyshev [C1], [C2] and Halphen [H]) to be related to the solvability of the Pell equation over polynomial rings. Here D is given in say $\mathbb{C}[X]$ and we ask whether there exist A, B also in $\mathbb{C}[X]$ with

$$A^2 - DB^2 = 1, \qquad B \neq 0. \tag{1.5}$$

A necessary condition is clearly that D has even degree. If the degree is 0 or 2, it is easy to see that the answer is always yes. If it is 4, then the answer is usually no. For example, introducing another parameter t (algebraically related to λ) we find that the answer is yes for $D = X^4 + X + t$ if and only if the point (0, 1) on the elliptic curve $y^2 = x^3 - 4tx + 1$ ($256t^3 \neq 27$) is torsion. In [MZ2, pp. 1677, 1678] we showed that if λ in $\mathbb{C}$ is such that just $(2, \sqrt{2(2-\lambda)})$ in (1.1) is torsion, then λ is in $\overline{\mathbb{Q}}$; and a similar argument holds for the t here.

But if D has degree 6, then we are in a situation analogous to the full (1.1): there is a point Π on the Jacobian such that $n\Pi = 0$ for some positive integer n. In this way we can handle one-parameter families. For the sake of illustration we shall restrict ourselves to the example $D = X^6 + X + t$, and we shall prove the following result.

Theorem P1. *There are at most finitely many complex t for which there exist A and $B \neq 0$ in $\mathbb{C}[X]$ with $A^2 - (X^6 + X + t)B^2 = 1$.*

There are some such t; for example with $t = 0$ we have

$$(2X^5 + 1)^2 - (X^6 + X)(2X^2)^2 = 1, \tag{1.6}$$

found quickly with continued fractions (see below). But we will show with the help of calculations over the finite fields $\mathbb{F}_3$ and $\mathbb{F}_5$ by Olaf Merkert that (1.5) is not solvable with $t = 1$ and $X^6 + X + 1$. In Theorem P2 below we will see that Theorem P1 is best possible in the natural sense that its analogue for $A^2 - (X^6 + X + t)B^2$ of degree at most one is false.

We can consider other one-parameter families of sextic D like $F(X)(X - t)$ for fixed quintic F in say $\mathbb{Q}[X]$, related to those considered by Ellenberg, Elsholtz, Hall and Kowalski in [EEHK] and [EHK]. But the example $D = X(X^2 + 1)(X^3 + X + t)$ has generic solution

$$(2t^{-1}X^3 + 2t^{-1}X + 1)^2 - D(2t^{-1})^2 = 1$$

and so a solution for all complex $t \neq 0$. There is a more complicated example for

$$D = X(X^5 - 10tX^4 + 35t^2X^3 - 50t^3X^2 + 25t^4X - t^{10} - 2t^5 - 1)$$

in McMullen's paper [Mc, pp. 665, 666]. Some deep results of Nadel [N] suggest that such identities are rare, and even for example that there might be an absolute upper bound on the degrees of A and B in (1.5) for any sextic over $\mathbb{C}(t)$. Nevertheless the identities show that at least one condition is needed to guarantee finiteness. This turns out to involve the points at infinity. And we will see that the more subtle example $D = X^6+X^2+t$ leads to solvability again for an infinite but countable set of t, as for $X^4 + X + t$; this gives an extra condition which turns out to involve the simplicity of the Jacobian. To check this, various methods are available; see especially the papers [K1], [K2] of Katz and the work [St] of Stoll quoted in the book [CF] of Cassels and Flynn.

If D is generically not square-free, such as $(X - t)^2(X^4 - 1)$ or more interestingly $X^2(X^4 + X + t)$, then the problem reduces to one about extensions of elliptic schemes by $\mathbb{G}_{\mathrm{m}}$, so the methods of [Bert3] and [BMPZ] are applicable (see also Section 3 of Schinzel's paper [Schi]).

The connection with integration of algebraic functions in elementary terms has also been classically known since Abel (and his functions) and Chebyshev (for elliptic functions, with his "pseudo-elliptic integrals"). In fact, our Theorem P1 for $D = X^6 + X + t$ is equivalent to the assertion that there are at most finitely many complex t for which there exists a non-zero E in $\mathbb{C}[X]$ of degree at most 4 such that $E/\sqrt{D}$ is integrable in elementary terms. As $D'/\sqrt{D}$ integrates to $2\sqrt{D}$, we cannot go up to degree 5 here. As above, the example (1.6) for $t = 0$ leads to

$$\int \frac{5X^2}{\sqrt{X^6 + X}}\, dX = \log\left(\frac{1}{2} + X^5 + X^2\sqrt{X^6 + X}\right). \tag{1.7}$$

It is interesting to compare this version of our Theorem P1 with one in the book [Dave] of Davenport. His Theorem 7 (p. 90) says that if an algebraic function $f(X, t)$ is not generically integrable in elementary terms, then there are at most finitely many complex t at which the specialised function is integrable in elementary terms. In fact, parts of his proof are unclear and we intend to investigate this more fully in future work. Here it will be necessary to go beyond semiabelian varieties.

Our Theorem P1 shows in particular that imitating the classical continued fraction algorithm for the Pell equation over $\mathbb{Z}$ will not work for $\mathbb{C}[X]$; a general fact also known since Abel and Chebyshev (see also the article [PT] of van der Poorten and Tran which also covers all the above connections, with illuminating examples). Through this link we deduce that there are at most finitely many complex t such that the continued fraction of

$$\sqrt{X^6 + X + t} = X^3 \sum_{k=0}^{\infty} \binom{1/2}{k} (X^{-5} + tX^{-6})^k = X^3 + \frac{1}{2}X^{-2} + \frac{t}{2}X^{-3} - \frac{1}{8}X^{-7} + \cdots$$

in the quotient field $\mathbb{C}((X^{-1}))$ of the ring of power series in X^{-1} is periodic. In the usual notation $[a_0; a_1, a_2, \ldots]$ it starts

$$a_0 = X^3, \quad a_1 = 2X^2 - 2tX + 2t^2, \quad a_2 = -\frac{1}{2t^3}X - \frac{1}{2t^2}, \quad a_3 = -8t^6X + 16t^7 \quad (t \neq 0).$$

But for $t = 0$ as in (1.6) and (1.7) we find

$$\sqrt{X^6 + X} = [X^3; 2X^2, 2X^3, 2X^2, 2X^3, 2X^2, 2X^3, 2X^2, 2X^3, \ldots]$$

with period 2.

When solving a Pell equation $a^2 - db^2 = 1$ over $\mathbb{Z}$, one notes that a/b must be a good rational approximation to $\sqrt{d}$. But constructing such good approximations by the Box Principle gives infinitely many solutions only of the equation $a^2 - db^2 = m$ for some fixed m, "almost the Pell equation". To obtain $m = 1$ an extra application of the Box Principle is needed.

Analogous considerations for general D in $\mathbb{C}[X]$ of even degree, such as the continued fraction algorithm or Padé approximation or linear algebra, will solve only

$$A^2 - DB^2 = M, \tag{1.8}$$

where for D of degree 6 the polynomial M (which cannot be prescribed in advance) has degree at most 2. Again for the sake of illustration we restrict ourselves to $D = X^6+X+t$; thus we get degree at most 0 for at most finitely many t. It is now natural to investigate the intermediate situation of degree at most 1. Here we have a generic example

$$(X^3)^2 - (X^6 + X + t)(1)^2 = -X - t \tag{1.9}$$

holding for all complex t. We take this into account first by proving

Theorem P2. *There are infinitely many complex t for which there exist A and non-constant B in $\mathbb{C}[X]$ and $c' \neq 0$, c in $\mathbb{C}$ with $A^2 - (X^6 + X + t)B^2 = c'X + c$.*

This situation corresponds to a point Π on the Jacobian of a curve such that $n\Pi$ lies on a fixed embedding of the curve, rather than $n\Pi = 0$ as for Theorem P1 above. In this sense Theorem P1 is best possible.

Then we show the set of t in Theorem P2 is countable provided we stay away from the generic example (1.9).

But this set seems more mysterious than that for $X^6 + X^2 + t$ (or $X^4 + X + t$). We have not even been able to prove that the above set is not the whole of $\overline{\mathbb{Q}}$! Suspecting a link with Chabauty's method for diophantine equations, we consulted Flynn, who very quickly did this and even showed for example that the set does not meet $7\mathbb{Z}$ except for $t = 0$. With his kind permission we include his proof as an Appendix to the present paper.

Let us say something about our own proofs. That of our Theorem follows the general strategy of [MZ1]–[MZ4] and [PZ], but several new issues arise. For example we can no longer express the periods in terms of hypergeometric functions, so we have to live with the period integrals. We have to study equations

$$\mathbf{z} = x\mathbf{f} + y\mathbf{g} + u\mathbf{k} + v\mathbf{l} \tag{1.10}$$

where $\mathbf{f}, \mathbf{g}, \mathbf{k}, \mathbf{l}$ are basis elements of the period lattice of $\mathcal{A}$ and $\mathbf{z}$ is an abelian logarithm. Our coefficients x, y, u, v are real and their locus S in $\mathbb{R}^4$ is subanalytic, of dimension at most 2 because a complex curve has real dimension 2. When $\mathbf{z}$ corresponds to a torsion

point, say of order dividing some n, then we get a rational point in $(1/n)\mathbb{Z}^4$ on S. The work of Pila [Pil] provides for any $\epsilon > 0$ an upper bound for their number, of order at most n^ϵ as n tends to infinity, provided we avoid connected semialgebraic curves inside S.

If $\mathcal{V}$ itself is contained in an abelian subscheme of $\mathcal{A}$ of positive codimension, there is nothing to prove. Otherwise we are able to show that there are no connected semialgebraic curves inside S. This follows from the algebraic independence of the two components of $\mathbf{z}$ over the field generated by the components of $\mathbf{f}, \mathbf{g}, \mathbf{k}, \mathbf{l}$ in (1.10). Here the remark of Bertrand mentioned in [MZ3] and [MZ4] is especially valuable in circumventing the question of dependence relations already holding between these components, which would depend for example on the type of complex multiplication of $\mathcal{A}$. In [MZ2] the analogous independence was proved with relatively simple arguments involving monodromy on just f and g so essentially $SL_2(\mathbb{Z})$. Extending these arguments in [MZ3] to f, g, z, w was a rather more complicated matter; we deduced the required independence from a result of Bertrand, and we also gave a self-contained proof involving $SL_4(\mathbb{Z})$. In [MZ4] we had to appeal to more general work of André [An] (see also Bertrand's paper [Bert1]); and this suffices here too.

We conclude the proof as in [MZ2]–[MZ4] by appealing to Silverman's Specialization Theorem [Si1]; however, now the new abelian situation requires a result of David [Davi] on degrees of torsion points of the corresponding fibre of $\mathcal{A}$. If this fibre is itself simple then we deduce by contrast that the number of rational points is of order at least n^δ for some $\delta > 0$. But the fibre could well be non-simple. Such obstacles did not arise in our earlier work. Perhaps this situation could be controlled with the help of conjectures (or even theorems) of André–Oort type. However, here we can avoid such problems by exploiting an escape clause in [Davi] arising from the "obstruction subgroups" in the transcendence method. We can then use some comparatively elementary estimates from the first author's work [MW1] with Wüstholz to reduce to a pair of elliptic curves, which can be handled as in [MZ2]–[MZ4] to get n^δ as well. Comparison of this lower bound with the above upper bound leads to an estimate for n which suffices to prove the Theorem.

Here is a brief section-by-section account of this paper.

In Section 2 we show how to reduce our Theorem to a Proposition involving the special case of a curve C in $\mathcal{A} = J_{\lambda\kappa\theta}$, the Jacobian of the hyperelliptic curve $H_{\lambda\kappa\theta}$ of genus 2 defined by

$$Y^2 = X(X-1)(X-\lambda)(X-\kappa)(X-\theta). \tag{1.11}$$

Then in Section 3 we recall the main result of [Pil] on subanalytic sets. Our own set is constructed from elliptic logarithms defined in Section 4. The relevant algebraic independence result is then proved in Section 5 (or Appendix A). This then leads in Section 6 to the non-existence of Pila's semialgebraic curves in our set. Then in Sections 7 and 8 we record the consequences of the work of David and Silverman for our purposes, and the proof of the Proposition is completed in Section 9.

In Section 10 we check the example (1.3) and prove Theorem P1, explaining in more detail the connections with integration and continued fractions. In Section 11 we prove Theorem P2, and finally in Section 12 we make some further remarks. The Appendix by Victor Flynn contains a proof of his results mentioned above.

2. Reduction to a hyperelliptic curve

We noted in [MZ3, Section 2] that the above Conjecture is isogeny invariant in the following sense. Let $\mathcal{S}, \mathcal{S}'$ be semiabelian schemes defined over varieties over $\mathbb{C}$ and suppose that there is an isogeny ι from $\mathcal{S}$ to $\mathcal{S}'$. Then the Conjecture for $\mathcal{S}'$ implies the Conjecture for $\mathcal{S}$. In fact the argument holds with $\mathbb{C}$ generalised to any algebraically closed field $\mathcal{K}$ of zero characteristic, and for possible later use we maintain this generality in the present short section.

Now every simple abelian surface is isogenous to the Jacobian of a curve of genus 2 (see for example [LB, p. 348]), and every such curve is well-known to be hyperelliptic. The latter can easily be put in the form $H_{\lambda\kappa\theta}$ of (1.11) above (here λ, κ, θ are sometimes called the *Rosenhain coordinates*). Thus we have an isogeny ι from the $\mathcal{S} = \mathcal{A}$ of our Theorem to some $\mathcal{S}' = J_{\lambda\kappa\theta}$ as above. We may think of points of the Jacobian as unordered pairs $\{P, Q\}$ of points $P = (X, Y),\ Q = (U, V)$ on $H_{\lambda\kappa\theta}$ corresponding to the divisor $(P)+(Q)-2(\infty)$, where ∞ is the unique point at infinity on the curve, together with the unordered pairs $\{P, \infty\}$ and the group origin $\{\infty, \infty\} = O$. Here all $\{(X, Y), (X, -Y)\}$ are identified with O. This can be compared with the analogous symbol in the book [CF] of Cassels and Flynn (p. 3); however they have a sextic polynomial on the left-hand side of (1.11).

Let $\mathcal{V}$ be a curve in $\mathcal{S}$. Then $\iota(\mathcal{V})$ in $J_{\lambda\kappa\theta}$ is a curve C in the affine space $\mathbb{A}^7$ with coordinates $X, Y, U, V, \lambda, \kappa, \theta$. We will regard it as being parametrised by $(\xi, \eta, \mu, \nu, \lambda, \kappa, \theta)$ with $\xi, \eta, \mu, \nu, \lambda, \kappa, \theta$ functions in $\mathcal{K}(C)$.

If the points $P = (\xi, \eta)$, $Q = (\mu, \nu)$ satisfy $n\{P, Q\} = O$ for some positive integer n, then the whole of $\iota(\mathcal{V})$ lies in the corresponding zero-dimensional abelian subscheme, so the Theorem is trivial for $\mathcal{S}'$. Thus we are entitled to assume $n\{P, Q\} \neq O$ for all such integers.

If λ, κ, θ are constant on C, then the base variety can be considered as a point and the Theorem for $\mathcal{S}'$ follows from Manin–Mumford as mentioned in the Introduction.

From all these considerations, we see that our Theorem for $\mathcal{A}$ is implied by the following statement.

Proposition. *Let C in $\mathbb{A}^7$ be a curve defined over $\overline{\mathbb{Q}}$ and parametrised by*

$$\mathbf{c} = (\xi, \eta, \mu, \nu, \lambda, \kappa, \theta)$$

in $\overline{\mathbb{Q}}(C)^7$, and suppose that the Jacobian $J_{\lambda\kappa\theta}$ of the curve $H_{\lambda\kappa\theta}$ of genus 2 is simple and non-constant. Suppose that the points

$$P = (\xi, \eta), \qquad Q = (\mu, \nu)$$

lie on $H_{\lambda\kappa\theta}$ and the point $\{P, Q\}$ is not identically torsion on $J_{\lambda\kappa\theta}$. Then there are at most finitely many points $\mathbf{c}$ in $C(\mathbb{C})$ such that for

$$P(\mathbf{c}) = (\xi(\mathbf{c}), \eta(\mathbf{c})), \quad Q(\mathbf{c}) = (\mu(\mathbf{c}), \nu(\mathbf{c}))$$

the point $\{P(\mathbf{c}), Q(\mathbf{c})\}$ is torsion on $J_{\lambda(\mathbf{c})\kappa(\mathbf{c})\theta(\mathbf{c})}$.

We note that the functions

$$\lambda, \lambda - 1, \kappa, \kappa - 1, \theta, \theta - 1, \lambda - \kappa, \kappa - \theta, \theta - \lambda \tag{2.1}$$

are all identically non-zero by our genus assumption. In fact we can also assume the same about

$$\xi, \xi - 1, \xi - \lambda, \xi - \kappa, \xi - \theta, \mu, \mu - 1, \mu - \lambda, \mu - \kappa, \mu - \theta. \tag{2.2}$$

If say $\xi = \lambda$ identically then $2\{P, Q\} = \{Q, Q\}$ (the function $X - \lambda$ then having divisor $2(P) - 2(\infty)$). This is not identically torsion either, and so by doubling the original point in the Proposition we end up with the new $\xi = \mu$. Now if say $\mu = 1$ identically then $\{Q, Q\} = O$, contradicting the fact that the original point is not identically torsion.

3. Rational points

In this section we record the basic result of Pila [Pil] that we shall use in the algebraic case. We recall from [MZ2, Section 3] that a *naive-m-subanalytic* subset of $\mathbb{R}^s$ is a finite union of $\theta(D)$, where each D is a closed ball in $\mathbb{R}^m$ and each θ is real analytic from an open neighbourhood of D to $\mathbb{R}^s$. We also refer there for the definition of S^{trans}.

Lemma 3.1. *Suppose S is a naive-2-subanalytic subset of $\mathbb{R}^s$. Then for any $\epsilon > 0$ there is a $c = c(S, \epsilon)$ with the following property. For each positive integer n there are at most cn^{ϵ} rational points of S^{trans} in $(1/n)\mathbb{Z}^s$.*

Proof. See [MZ2, Lemma 2.1, p. 1680].

4. Functions

We will construct our naive-2-subanalytic subset S by means of the following functions. With λ, κ, θ in $\mathbb{C}(C)$ as in the Proposition and X, Y as in (1.11), we consider the standard integrals

$$\left(\int \frac{dX}{Y}, \int \frac{X\,dX}{Y} \right) \tag{4.1}$$

over loops. By the remark about (2.1) the set of $\mathbf{c}$ in $C(\mathbb{C})$ not satisfying

$$\lambda(\mathbf{c}) \neq 0, 1, \infty,\ \kappa(\mathbf{c}) \neq 0, 1, \infty,\ \theta(\mathbf{c}) \neq 0, 1, \infty,\ \ \lambda(\mathbf{c}) \neq \kappa(\mathbf{c}) \neq \theta(\mathbf{c}) \neq \lambda(\mathbf{c}) \tag{4.2}$$

is at most finite. We pick any $\mathbf{c}_*$ satisfying (4.2) and then pick four loops on $H_{\lambda(\mathbf{c}_*)\kappa(\mathbf{c}_*)\theta(\mathbf{c}_*)}$ generating the homology. These by (4.1) define functions $\mathbf{f}, \mathbf{g}, \mathbf{k}, \mathbf{l}$ to $\mathbb{C}^2$ at $\mathbf{c}_*$. We may extend them, at least locally, to the set of all $\mathbf{c}$ in C with (4.2), and they are analytic in $\lambda = \lambda(\mathbf{c})$, $\kappa = \kappa(\mathbf{c})$, $\theta = \theta(\mathbf{c})$. It is well-known that they are basis elements of a period lattice of $J_{\lambda\kappa\theta}$ with respect to $\left(\frac{dX}{Y}, \frac{X\,dX}{Y}\right)$. In particular, if we write $\exp_{\lambda\kappa\theta}$ for the associated exponential function from $\mathbb{C}^2$ to $J_{\lambda\kappa\theta}(\mathbb{C})$, we have

$$\exp_{\lambda\kappa\theta}(\mathbf{f}) = \exp_{\lambda\kappa\theta}(\mathbf{g}) = \exp_{\lambda\kappa\theta}(\mathbf{k}) = \exp_{\lambda\kappa\theta}(\mathbf{l}) = O.$$

Next let $P = (\xi, \eta),\ Q = (\mu, \nu)$ be as in the Proposition with ξ, η, μ, ν in $\mathbb{C}(C)$. We would like to define

$$\mathbf{z} = \left(\int_\infty^P \frac{dX}{Y} + \int_\infty^Q \frac{dX}{Y},\ \int_\infty^P \frac{X\,dX}{Y} + \int_\infty^Q \frac{X\,dX}{Y} \right) \tag{4.3}$$

as an abelian logarithm of $\{P, Q\}$ which is analytic in a suitable sense. This is also possible everywhere locally apart from finitely many exceptional points. In fact the remarks about (2.2) together with the discussion in [MZ3, Section 4], which replaces the curve integral with an X-integral, lead without difficulty to the following.

Write $\hat{C}$ for the set of points $\mathbf{c}$ of $C(\mathbb{C})$ with (4.2) and

$$\xi(\mathbf{c}),\ \mu(\mathbf{c}) \neq 0, 1, \infty,\ \lambda(\mathbf{c}),\ \kappa(\mathbf{c}),\ \theta(\mathbf{c})$$

as in (2.2). The points not in $\hat{C}$ still form at most a finite set. Then for any $\mathbf{c}_*$ in $\hat{C}$ and any sufficiently near $\mathbf{c}$ in $\hat{C}$ we can express the first component of $\mathbf{z}$ in (4.3) as a quadruple power series in

$$\lambda(\mathbf{c}) - \lambda(\mathbf{c}_*),\ \kappa(\mathbf{c}) - \kappa(\mathbf{c}_*),\ \theta(\mathbf{c}) - \theta(\mathbf{c}_*),\ \xi(\mathbf{c}) - \xi(\mathbf{c}_*),$$

and the second component as a quadruple power series in

$$\lambda(\mathbf{c}) - \lambda(\mathbf{c}_*),\ \kappa(\mathbf{c}) - \kappa(\mathbf{c}_*),\ \theta(\mathbf{c}) - \theta(\mathbf{c}_*),\ \mu(\mathbf{c}) - \mu(\mathbf{c}_*).$$

Also

$$\exp_{\lambda\kappa\theta}(\mathbf{z}) = \{P, Q\}. \tag{4.4}$$

5. Algebraic independence

In this section we consider the point $\mathbf{c}_*$ of $\hat{C}$ to be fixed. Then $\mathbf{f}, \mathbf{g}, \mathbf{k}, \mathbf{l}, \mathbf{z}$ are well-defined on a small neighbourhood N_* of $\mathbf{c}_*$. In order to prove $S^{\text{trans}} = S$ we will need the following result.

Lemma 5.1. *The coordinates of* $\mathbf{z}$ *are algebraically independent over* $\mathbb{C}(\mathbf{f}, \mathbf{g}, \mathbf{k}, \mathbf{l})$ *on* N_*.

Proof. This follows from [An, Theorem 3, p. 16], which actually specifies the transcendence degree of $K(\mathbf{z}, \tilde{\mathbf{z}})$ over $K = \mathbb{C}(C)(\mathbf{f}, \mathbf{g}, \mathbf{k}, \mathbf{l}, \tilde{\mathbf{f}}, \tilde{\mathbf{g}}, \tilde{\mathbf{k}}, \tilde{\mathbf{l}})$, where the extra functions are the corresponding integrals of the second kind with respect to say $\frac{X^2\,dX}{Y}, \frac{X^3\,dX}{Y}$. It is the dimension of the $\tilde{U}$ appearing in [An, Proposition 1, p. 5], or at least its relative counterpart in the context of [An, Section 4]. The E there is $J_{\lambda\kappa\theta}$ over C, for which our simplicity hypothesis implies that the only non-zero proper connected algebraic subgroup is O. And because $\{P, Q\}$ is not identically torsion, the E' there is also E, with rational homology isomorphic to $\mathbb{Q}^4$. Further because of simplicity the F there is a division algebra. And $u(\mathcal{X})$ there is from $\mathbb{Z}$ to $\mathbb{Z}\{P, Q\}$. So $F.u(\mathcal{X})$ is isomorphic to F. Thus we find dimension 4; and the present lemma follows on throwing away all the extra functions. See also Bertrand's article [Bert1, end of Section 4, p. 2786] as well as [Bert2, Theorem 4.3, p. 16].

6. A naive-2-subanalytic set

We now describe our naive-2-subanalytic subset S. First we construct local functions from C to $\mathbb{R}^4$. Recall that $\hat{C}$ is obtained from $C(\mathbb{C})$ by the removal of at most a finite set of points. Fix $\mathbf{c}_*$ in $\hat{C}$, choose $\mathbf{c}$ in $\hat{C}$ and then a path from $\mathbf{c}_*$ to $\mathbf{c}$ lying in $\hat{C}$. We can continue $\mathbf{f}, \mathbf{g}, \mathbf{k}, \mathbf{l}$ taking care to keep a homology basis.

The continuation of the functions z, w is a bit more troublesome, and it is convenient to also remove the singular points of C. Let C_0 be the finite subset which we have removed so far, and write $\widehat{C}$ for what remains. We can then speak of functions analytic on $\widehat{C}$. Now the discussion in [MZ3, Section 6], with $\exp_{\lambda\kappa\theta}$ instead of $\exp_\lambda$ and $\mathbf{z}_2 = x\mathbf{f} + y\mathbf{g} + u\mathbf{k} + v\mathbf{l} + \mathbf{z}_1$ instead of $z_2 = xf + yg + z_1$, shows that we can continue the function $(\mathbf{f}, \mathbf{g}, \mathbf{k}, \mathbf{l}, \mathbf{z})$ from a small neighbourhood of $\mathbf{c}_*$ to a small neighbourhood $N_\mathbf{c}$ of $\mathbf{c}$ in $\widehat{C}$. The end result is a function $(\mathbf{f}_\mathbf{c}, \mathbf{g}_\mathbf{c}, \mathbf{k}_\mathbf{c}, \mathbf{l}_\mathbf{c}, \mathbf{z}_\mathbf{c})$ analytic on $N_\mathbf{c}$. Write $\Omega_\mathbf{c}$ for the period lattice of $J_{\lambda(\mathbf{c})\kappa(\mathbf{c})\theta(\mathbf{c})}$ with respect to $\left(\frac{dX}{Y}, \frac{X\,dX}{Y}\right)$.

Lemma 6.1. *The coordinates of $\mathbf{z}_\mathbf{c}$ are algebraically independent over $\mathbb{C}(\mathbf{f}_\mathbf{c}, \mathbf{g}_\mathbf{c}, \mathbf{k}_\mathbf{c}, \mathbf{l}_\mathbf{c})$ on $N_\mathbf{c}$. Further $\Omega_\mathbf{c} = \mathbb{Z} f_c + Z g_c + Z k_c + Z l_c$ on $N_\mathbf{c}$.*

Proof. We could continue an algebraic dependence relation backwards to get the same relation between $\mathbf{f}, \mathbf{g}, \mathbf{k}, \mathbf{l}, \mathbf{z}$ on a neighbourhood of $\mathbf{c}_*$; however, this would contradict Lemma 5.1. The assertions about $\Omega_\mathbf{c}$ follow because we kept a homology basis during the continuation.

It follows that we can define $x_\mathbf{c}, y_\mathbf{c}, u_\mathbf{c}, v_\mathbf{c}$ on $N_\mathbf{c}$ by the equation

$$\mathbf{z}_\mathbf{c} = x_\mathbf{c}\mathbf{f}_\mathbf{c} + y_\mathbf{c}\mathbf{g}_\mathbf{c} + u_\mathbf{c}\mathbf{k}_\mathbf{c} + v_\mathbf{c}\mathbf{l}_\mathbf{c} \tag{6.1}$$

and its complex conjugate

$$\overline{\mathbf{z}_\mathbf{c}} = x_\mathbf{c}\overline{\mathbf{f}_\mathbf{c}} + y_\mathbf{c}\overline{\mathbf{g}_\mathbf{c}} + u_\mathbf{c}\overline{\mathbf{k}_\mathbf{c}} + v_\mathbf{c}\overline{\mathbf{l}_\mathbf{c}}$$

so that $x_\mathbf{c}, y_\mathbf{c}, u_\mathbf{c}, v_\mathbf{c}$ are real-valued.

Now we can define S. We use the standard maximum norm on $\mathbb{C}^7$. For small $\delta > 0$ (to be specified later) we define C^δ as the set of $\mathbf{c}$ in C satisfying $|\mathbf{c}| \le 1/\delta$ and

$$|\mathbf{c} - \mathbf{c}_0| \ge \delta$$

for each $\mathbf{c}_0$ in the finite set C_0.

Shrinking $N_\mathbf{c}$ if necessary, we can choose a local analytic isomorphism $\varphi_\mathbf{c}$ from $N_\mathbf{c}$ to an open subset of $\mathbb{C}$ (i.e. $\mathbb{R}^2$). Choose any closed disc $D_\mathbf{c}$ inside $\varphi_\mathbf{c}(N_\mathbf{c})$ centred at $\mathbf{c}$, and define

$$\theta_\mathbf{c} = (x_\mathbf{c}, y_\mathbf{c}, u_\mathbf{c}, v_\mathbf{c}) \circ \varphi_\mathbf{c}^{-1}$$

from $D_\mathbf{c}$ to $\mathbb{R}^4$. By compactness there is a finite set Π of $\mathbf{c}$ such that the $\varphi_\mathbf{c}^{-1}(D_\mathbf{c})$ cover C^δ. Then our naive-2-subanalytic subset $S = S^\delta$ in $\mathbb{R}^4$ is defined as the union of all $\theta_\mathbf{c}(D_\mathbf{c})$ over $\mathbf{c} \in \Pi$.

Lemma 6.2. *We have $S^{\mathrm{trans}} = S$.*

Proof. Because every semialgebraic surface contains semialgebraic curves, it will suffice to deduce a contradiction from the existence of a semialgebraic curve B_s lying in S. Now B_s is Zariski-dense in its Zariski-closure B, a real algebraic curve. Thus we can find a subset $\hat{B}$ of B, also Zariski-dense in B, contained in some $\theta_{\mathbf{c}}(D_{\mathbf{c}})$. It will suffice to know that $\hat{B}$ is infinite. Then $\hat{B} = \theta_{\mathbf{c}}(E)$ for some infinite subset E of $D_{\mathbf{c}}$.

Now (6.1) shows that the components of $\mathbf{z_c}$ lie in $\Phi = \mathbb{C}(x_{\mathbf{c}}, y_{\mathbf{c}}, u_{\mathbf{c}}, v_{\mathbf{c}}, \mathbf{f_c}, \mathbf{g_c}, \mathbf{k_c}, \mathbf{l_c})$. But if we restrict to $\varphi_{\mathbf{c}}^{-1}(E)$, then Φ has transcendence degree at most 1 over $\mathbb{C}(\mathbf{f_c}, \mathbf{g_c}, \mathbf{k_c}, \mathbf{l_c})$. It follows that the components of $\mathbf{z_c}$ are algebraically dependent over $\mathbb{C}(\mathbf{f_c}, \mathbf{g_c}, \mathbf{k_c}, \mathbf{l_c})$ on $\varphi_{\mathbf{c}}^{-1}(E)$. More precisely, with independent variables $\mathbf{T_f}, \mathbf{T_g}, \mathbf{T_k}, \mathbf{T_l}, \mathbf{T_z}$, there exists a polynomial A in $\mathbb{C}[\mathbf{T_f}, \mathbf{T_g}, \mathbf{T_k}, \mathbf{T_l}, \mathbf{T_z}]$ such that the relation $A(\mathbf{f_c}, \mathbf{g_c}, \mathbf{k_c}, \mathbf{l_c}, \mathbf{z_c}) = 0$ holds on $\varphi_{\mathbf{c}}^{-1}(E)$ and $A(\mathbf{f_c}, \mathbf{g_c}, \mathbf{k_c}, \mathbf{l_c}, \mathbf{T_z})$ is not identically zero in $\mathbb{C}(\mathbf{f_c}, \mathbf{g_c}, \mathbf{k_c}, \mathbf{l_c})[\mathbf{T_z}]$. By a standard principle for analytic functions ("Identity Theorem" or [L, p. 85]) this relation persists on all of $N_{\mathbf{c}}$. And now we have a contradiction with Lemma 6.1. Thus the present lemma is proved.

We are all set up for an efficient application of Lemma 3.1. It will turn out that every $\mathbf{c}$ in our Proposition leads to many rational points on S, and of course we have to estimate their denominator. This we do in the next short section.

7. Orders of torsion

We use the standard absolute Weil height

$$h(\alpha) = \frac{1}{[\mathbb{Q}(\alpha) : \mathbb{Q}]} \sum_v \log \max\{1, |\alpha|_v\}$$

of an algebraic number α, where v runs over a suitably normalized set of valuations; and also the standard extension to vectors using the maximum norm. See for example [Si2, p. 208].

Lemma 7.1. *There is a constant $c = c(C)$ with the following property. Suppose that for some $\mathbf{a}$ in $\hat{C}$ the point $\{P(\mathbf{a}), Q(\mathbf{a})\}$ on $J_{\lambda(\mathbf{a})\kappa(\mathbf{a})\theta(\mathbf{a})}$ has finite order n. Then $\mathbf{a}$ is algebraic, and*

$$n \le c[\mathbb{Q}(\mathbf{a}) : \mathbb{Q}]^7 (1 + h(\mathbf{a}))^6.$$

Proof. It is clear that $\mathbf{a}$ is algebraic, otherwise $\{P, Q\}$ would be identically torsion on C, contradicting a hypothesis of the Proposition.

As for the upper bound, if the (principally polarized) $J = J_{\lambda(\mathbf{a})\kappa(\mathbf{a})\theta(\mathbf{a})}$ is simple, then a result of David [Davi, Théorème 1.2, p. 121] gives

$$\frac{n^{1/2}}{\log n} \le c_1 d_\Pi (d_J^{3/2} \log d_J) \max\{1, h\}^{3/2}$$

where d_J is the degree of a field of definition k of J over $\mathbb{Q}$, d_Π is the degree of a field of definition of $\{P(\mathbf{a}), Q(\mathbf{a})\}$ over k, h is the semistable Faltings height of J, and c_1 is an absolute constant. We can take $k = \mathbb{Q}(\lambda(\mathbf{a}), \kappa(\mathbf{a}), \theta(\mathbf{a}))$ and so $d_J \le c_2 \mathcal{D}$ for

$\mathcal{D} = [\mathbb{Q}(\mathbf{a}) : \mathbb{Q}]$ with c_2 independent of $\mathbf{a}$. Also since λ, κ, θ are not all constant, if for example λ is not constant, then each of the coordinates ξ, η, μ, ν of P and Q is algebraic over $\mathbb{Q}(\lambda)$. Thus at $\mathbf{c} = \mathbf{a}$ we deduce $d_\Pi \le c_2$. And then $h \le c_3(1 + h(\mathbf{a}))$ by well-known properties of the Faltings height (see for example the discussion [Davi, p. 123]). The required result follows, with slightly smaller exponents.

But what if J is not simple? It may then be that certain conjectures of André–Oort type lead anyway to at most finitely many possibilities for $\mathbf{a}$, as required in our original Proposition. But in the absence of proofs we can reduce to an elliptic situation as follows.

Our J, being the Jacobian of a curve of genus 2, can be embedded in projective $\mathbb{P}_{15}$; see for example [CF, p. 8] after applying a fractional linear transformation to replace the quintic by a sextic. This is more or less the same embedding that David uses for $A(\tau)$ in his work (but the quintic itself gives embeddings in $\mathbb{P}_8$—see for example Grant [G, p. 101]). Thus consulting [Davi, equation (28), p. 156] we find an algebraic subgroup $B \ne J$ of J. In fact Philippon's multiplicity estimate used there (p. 159) guarantees that B is connected. If $B = 0$ then we can deduce equation (29) of [Davi, p. 156] and this leads to a much better bound, say $n \le c[\mathbb{Q}(\mathbf{a}) : \mathbb{Q}]^4(1 + h(\mathbf{a}))^3$. So it remains only to treat the case that B is an elliptic curve. We note by [MW1, Lemma 2.2, p. 414] that B is defined over an extension of k of degree at most 3^{256}. And we get the estimate $T\Delta L \le c_4(LN^2)^2$ for the degree Δ of B in the embedding, where T, L, N are defined in [Davi, p. 152] and again c_4 is absolute. We find $\Delta \le c_5(\mathcal{D}\log\mathcal{D})^2h^2$ for absolute c_5.

We can now apply [MW1, Lemma 1.4, p. 413] to find another elliptic curve B' in J (so also defined over an extension of k of degree at most 3^{256}) together with an isogeny ϕ from $B \times B'$ to J of degree at most Δ^2. The dual isogeny ψ from J to $B \times B'$ has degree at most Δ^6. Thus by standard properties of Faltings heights we have

$$h(B) \le h(B \times B') + c_7 \le h + \tfrac{1}{2}\log(\Delta^6) + c_7 \le c_8(1 + \log\mathcal{D} + h(\mathbf{a})) \tag{7.1}$$

with the same bound for $h(B')$. We can reduce B to Weierstrass form E without increasing the field of definition; also $h(E) = h(B)$. Now the argument of [MZ3, Lemma 7.1] shows that the order m of the projection of $\psi(\{P(\mathbf{a}), Q(\mathbf{a})\})$ on E through B satisfies

$$m \le c_8(\mathcal{D}\max\{1, h(j_E)\} + \mathcal{D}\log\mathcal{D})$$

for the corresponding j-invariant. It is well-known that $h(j_E)$ is of the same order of magnitude as $h(E)$, so appealing to (7.1) we deduce $m \le c_9\mathcal{D}(1 + h(\mathbf{a}) + \log\mathcal{D})$.

We get the same bound for the order m' of the projection of $\psi(\{P(\mathbf{a}), Q(\mathbf{a})\})$ on B'. Thus $\psi(\{P(\mathbf{a}), Q(\mathbf{a})\})$ has order at most mm'. Applying ϕ back shows finally that $n \le \Delta^2 mm'$; and putting everything together gives the required result.

We could use more directly the factorisation estimates of [MW2] to get B, B' and ϕ, but the exponents involved would be astronomical.

8. Heights

In view of the following result we can eliminate the height dependence in Lemma 7.1.

Lemma 8.1. *There is a constant $c = c(C)$ with the following property. Suppose that for some $\mathbf{a}$ in $\hat{C}$ the point $\{P(\mathbf{a}), Q(\mathbf{a})\}$ has finite order. Then $h(\mathbf{a}) \le c$.*

Proof. This is a consequence of Silverman's Specialization Theorem [Si1, p. 197], because $\{P(\mathbf{c}), Q(\mathbf{c})\}$ is not identically of finite order; note that our family of abelian surfaces has no non-trivial constant part because it is generically simple.

Another advantage of bounded height is the following easy remark, already found in [MZ4], concerning the sets C_0 and C^δ of Section 6.

Lemma 8.2. *Let K be a number field containing the coordinates of the points of C_0. For any constant c there is a positive $\delta = \delta(C, K, c)$ depending only on C, K and c with the following property. Suppose $\mathbf{a}$ is algebraic on C, not in C_0, with $h(\mathbf{a}) \le c$. Then there are at least $\frac{1}{2}[K(\mathbf{a}) : K]$ conjugates of $\mathbf{a}$ over K lying in C^δ.*

Proof. See [MZ4, Lemma 8.2].

9. Proof of Proposition

We will need the following result from [MZ3].

Lemma 9.1. *Suppose $f_0, f_1, \dots, f_s$ are analytic in an open neighbourhood N of a compact set $\mathcal{K}$ in $\mathbb{C}$ and f_0 is linearly independent of $f_1, \dots, f_s$ over $\mathbb{C}$. Then there is $c = c(f_0, f_1, \dots, f_s)$ with the following property. For any complex numbers $a_1, \dots, a_s$ the function $F = f_0 + a_1 f_1 + \cdots + a_s f_s$ has at most c different zeros on $\mathcal{K}$.*

Proof. See [MZ3, Lemma 9.1, p. 463].

To prove our Proposition we fix any positive $\epsilon < 1/7$. We use $c, c_1, c_2, \dots$ for positive constants depending only on C. We have to show that there are at most finitely many $\mathbf{a}$ such that $\Pi(\mathbf{a}) = \{P(\mathbf{a}), Q(\mathbf{a})\}$ has finite order on $J_{\lambda(\mathbf{a})\kappa(\mathbf{a})\theta(\mathbf{a})}$. By Lemma 7.1 each such $\mathbf{a}$ is algebraic, say of degree $\mathcal{D} = [\mathbb{Q}(\mathbf{a}) : \mathbb{Q}]$, and thanks to Lemma 8.1 and the Northcott property it will suffice to prove that $\mathcal{D} \le c$. We will actually argue with a single $\mathbf{a}$.

Next, Lemma 7.1 together with Lemma 8.1 shows that there is a positive integer

$$n \le c_1 \mathcal{D}^7 \tag{9.1}$$

such that

$$n\Pi(\mathbf{a}) = O. \tag{9.2}$$

Fix a number field K containing a field of definition for the curve C. By Lemmas 8.1 and 8.2 the algebraic $\mathbf{a}$ has at least $\frac{1}{2}[K(\mathbf{a}) : K]$ conjugates over K in some C^δ; here $\delta = c_2$. Now C^δ is contained in the union of at most c_3 closed sets $\varphi_{\mathbf{c}}^{-1}(D_{\mathbf{c}})$, and so there is $\mathbf{c}$ such that $\varphi_{\mathbf{c}}^{-1}(D_{\mathbf{c}})$ contains at least $c_4[K(\mathbf{a}) : K]$ conjugates $\sigma(\mathbf{a})$. And the corresponding conjugate point $\sigma(\Pi(\mathbf{a})) = \Pi(\sigma(\mathbf{a}))$ also satisfies $n\Pi(\sigma(\mathbf{a})) = O$.

We claim that each point $\Theta_\sigma = \theta_{\mathbf{c}}(\varphi_{\mathbf{c}}(\sigma(\mathbf{a})))$ lies in $\mathbb{Q}^4$ and even that $n\Theta_\sigma$ lies in $\mathbb{Z}^4$.

Now the function $\theta_{\mathbf{c}}$ arises from continuations $\mathbf{f_c}, \mathbf{g_c}, \mathbf{k_c}, \mathbf{l_c}, \mathbf{z_c}$ of the functions in Section 6. We deduce from (4.4) that

$$\exp_{\lambda(\mathbf{c})\kappa(\mathbf{c})\theta(\mathbf{c})}(\mathbf{z_c}) = \{P(\mathbf{c}),\, Q(\mathbf{c})\} \tag{9.3}$$

on $N_{\mathbf{c}}$. At $\sigma(\mathbf{a})$ this implies

$$\exp_{\lambda(\sigma(\mathbf{a}))\kappa(\sigma(\mathbf{a}))\theta(\sigma(\mathbf{a}))}(n\mathbf{z_c}(\sigma(\mathbf{a}))) = O. \tag{9.4}$$

It follows that $n\mathbf{z_c}(\sigma(\mathbf{a}))$ lies in the period lattice $\Omega_{\lambda(\sigma(\mathbf{a}))\kappa(\sigma(\mathbf{a}))\theta(\sigma(\mathbf{a}))}$, which by Lemma 6.1 is just $\mathbb{Z}\mathbf{f_c}(\sigma(\mathbf{a})) + \mathbb{Z}\mathbf{g_c}(\sigma(\mathbf{a})) + \mathbb{Z}\mathbf{k_c}(\sigma(\mathbf{a})) + \mathbb{Z}\mathbf{l_c}(\sigma(\mathbf{a}))$. Thus (6.1) shows that

$$nx_{\mathbf{c}}(\sigma(\mathbf{a})),\ ny_{\mathbf{c}}(\sigma(\mathbf{a})),\ nu_{\mathbf{c}}(\sigma(\mathbf{a})),\ nv_{\mathbf{c}}(\sigma(\mathbf{a}))$$

lie in $\mathbb{Z}$. Hence indeed $n\Theta_\sigma$ lies in $\mathbb{Z}^4$ as claimed.

So now each Θ_σ in the set S of Section 6 has common denominator dividing n. By Lemmas 3.1 and 6.2, the number of such values Θ_σ is at most $c_5 n^\epsilon$. By (9.1) this is at most $c_6\mathcal{D}^{7\epsilon}$. Let $\Theta = (x, y, u, v)$ be one of these values. For any σ with $\theta_{\mathbf{c}}(\varphi_{\mathbf{c}}(\sigma(\mathbf{a}))) = \Theta$ the expression $\mathbf{z_c}(\sigma(\mathbf{a}))$ is

$$x_{\mathbf{c}}(\sigma(\mathbf{a}))\mathbf{f_c}(\sigma(\mathbf{a})) + y_{\mathbf{c}}(\sigma(\mathbf{a}))\mathbf{g_c}(\sigma(\mathbf{a})) + u_{\mathbf{c}}(\sigma(\mathbf{a}))\mathbf{k_c}(\sigma(\mathbf{a})) + v_{\mathbf{c}}(\sigma(\mathbf{a}))\mathbf{l_c}(\sigma(\mathbf{a})),$$

which is

$$x\mathbf{f_c}(\sigma(\mathbf{a})) + y\mathbf{g_c}(\sigma(\mathbf{a})) + u\mathbf{k_c}(\sigma(\mathbf{a})) + v\mathbf{l_c}(\sigma(\mathbf{a})).$$

Lemma 6.1 implies that for example the first coordinate of $\mathbf{z_c}$ is linearly independent of the first coordinates of $\mathbf{f_c}, \mathbf{g_c}, \mathbf{k_c}, \mathbf{l_c}$. So Lemma 9.1 shows that the number of σ for each Θ is at most c_7.

Thus the total number of $\sigma(\mathbf{a})$ is at most $c_8\mathcal{D}^{7\epsilon}$. Now this contradicts the lower bound $c_4\mathcal{D}$ noted just after (9.2), provided $\mathcal{D}$ is sufficiently large. As observed near the beginning of this section, that suffices to prove our Proposition.

10. Examples and the Pell equation

It was shown in [M3, p. 294] that the Jacobian of (1.4) is identically simple (and even that the endomorphism ring is $\mathbb{Z}$). It has good reduction at all the points (1.2). By the equivalence of (a),(b) in Theorem 1 of Serre–Tate [ST, p. 493] any torsion point is unramified outside (1.2). However the point arising from (1.3) is ramified for example at $\lambda = 2$, as this is already true of the bisymmetric function

$$\sqrt{2(2-\lambda)(2-\lambda^2)(2-\lambda^4)}\sqrt{6(3-\lambda)(3-\lambda^2)(3-\lambda^4)}.$$

Thus the point is not identically torsion and our result applies.

To deal with the Pell equation $A^2 - DB^2 = 1$ with squarefree D of degree 6 we choose any field K (not of characteristic 2) over which D is defined, and we consider the hyperelliptic curve H_D defined by $Y^2 = D(X)$. This is singular at infinity with two points ∞^+, ∞^- on a non-singular model; we may fix them by stipulating that the function

$X^3 \pm Y$ has a zero at $\infty^{\pm}$. We pass to a non-singular model in the standard way by selecting any three zeros of D, and finding the fractional linear transformation taking them to $0, 1, \infty$. With λ, κ, θ as the other three images this gives a birational map β from H_D to $H_{\lambda\kappa\theta}$. Of course this might no longer be defined over K, but certainly over a splitting field of D. We may then speak of $\beta(\infty^{\pm})$ as points on $H_{\lambda\kappa\theta}$ (whose Jacobian is $J_{\lambda\kappa\theta}$ as in the discussion of Section 2). We now record the following fairly well-known result.

Lemma 10.1. (i) *Suppose there are A, B in $K[X]$ and $c \neq 0$ in K with $A \neq 0$ of degree d such that $A^2 - DB^2 = c$. Then for $n = \pm d$ the function $A(X) + YB(X)$ on H_D has a zero of order n at ∞^+ and a pole of order n at ∞^-, and no other zeros or poles.*

(ii) *Suppose there are A, B in $K(X)$ with $A \neq 0$ and n such that $A(X) + YB(X)$ on H_D has a zero of order n at ∞^+ and a pole of order n at ∞^-, and no other zeros or poles. Then $n = \pm d$ and A, B are in $K[X]$ with A of degree d such that $A^2 - DB^2 = c$ for some $c \neq 0$ in K.*

(iii) *Suppose D splits completely over K. Then there are A, B in $K[X]$ with $B \neq 0$ and $c \neq 0$ in K such that $A^2 - DB^2 = c$ if and only if the point $\{\beta(\infty^+), \beta(\infty^+)\}$ is torsion on $J_{\lambda\kappa\theta}$.*

Proof. In the situation of (i) we have $f^+ f^- = c$ for the functions $f^+ = A(X) + YB(X)$, $f^- = A(X) - YB(X)$ on H_D. So the only possible zeros and poles are at ∞^+, ∞^-. Since the number of zeros equals the number of poles, there is an integer n such that f^+ has a zero of order n at ∞^+ and a pole of order n at ∞^-. Now $f^+ + f^- = 2A$ has poles of order $|n|$ at ∞^+, ∞^- and no other poles. Thus $|n| = d$, and this proves (i).

In the situation of (ii) with $f^+ = A(X) + YB(X)$ we deduce that $f^- = A(X) - YB(X)$ has a pole of order n at ∞^+ and a zero of order n at ∞^-. Thus $f^+ f^-$ is a constant c, non-zero because $A \neq 0$. Also $f^+ + f^- = 2A$ has no poles at finite points, so it must be a polynomial. And finally because D is squarefree we see from $DB^2 = A^2 - c$ that B is also a polynomial. This brings us back to (i) and thereby completes the proof of (ii).

Finally in the situation of (iii) the existence of A, B with $B \neq 0$ implies A is not constant, so this gives from (i) a function $f^+ = A(X) + YB(X)$ from which we derive a function g on $H_{\lambda\kappa\theta}$ with a zero of order n at $P^+ = \beta(\infty^+)$ and a pole of order n at $P^- = \beta(\infty^-)$. Here $d \geq 1$ so $n \neq 0$. If $P^+ = (x, y)$ then $P^- = (x, -y)$ so there is an obvious linear function l with simple zeros at P^+, P^- and a double pole at ∞ on $H_{\lambda\kappa\theta}$. So looking at gl^n shows that $n\{P^+, P^+\} = O$. And conversely if $n\{P^+, P^+\} = O$ for say $n \geq 1$ then going backwards we find a function on H_D with a zero of order n at ∞^+ and a pole of order n at ∞^-. This can be written as $f^+ = A(X) + YB(X) \neq 0$ for A, B in $K(X)$; here both $A \neq 0$ and $B \neq 0$ otherwise f^+ could have no genuine zero at ∞^+, and so we are back to (ii). This completes the proof of the present lemma.

Of course if K is algebraically closed then for any $c \neq 0$ the solvability of

$$A^2 - DB^2 = c, \quad B \neq 0 \tag{10.1}$$

is equivalent to the solvability of Pell (1.5). But actually this holds for any K, because (10.1) implies $A \neq 0$ and then $A_1^2 - DB_1^2 = 1$ for $A_1 = c^{-1}(A^2 + DB^2)$ and $B_1 = 2c^{-1}AB \neq 0$.

We already noted in Section 1 that the solvability is equivalent to the periodicity of the functional continued fraction of $\sqrt{D}$; this was observed by Abel [Ab], where also the restriction to degree 6 is not essential. See also Chapter XIV of [H] as well as [PT] and the works [AR] of Adams and Razar, [Berr] of Berry, [P] of Paysant-Le Roux, [Schi] of Schinzel and [Schm] of Wolfgang Schmidt (however, we could not consult the paper [HL] of Hellegouarch and Lozach). See also [BC] of Bombieri and Cohen for connections with the arithmetic behaviour of Padé approximants.

And it is also equivalent to the existence of a non-zero polynomial E of degree at most 4 such that $E/\sqrt{D}$ is integrable in elementary terms. In that case, E must have degree 2, and it must be proportional to A'/B, and we have

$$\int \frac{E(X)}{\sqrt{D(X)}}\,dX = \log(A(X) + B(X)\sqrt{D(X)}). \tag{10.2}$$

This also *mutatis mutandis* is not restricted to degree 6; see [Ab], [C1], [C2], [H] and again [PT].

Now we prove Theorem P1 for $D = X^6 + X + t$. We start by showing that there are no $A, B \neq 0$ in $K[X]$ with $A^2 - DB^2 = 1$ for $K = \overline{\mathbb{C}(t)}$ (see also [Za, Remark 3.4.2, p. 85]).

Otherwise taking conjugates of the $f^+ = A + YB$ of Lemma 10.1(i) over $\mathbb{C}(t)$ (over which H_D, ∞^+, ∞^- are defined) would give $(f^+)^\sigma = A^\sigma + YB^\sigma$ with the same divisor. So $(f^+)^\sigma = cf^+$ for some non-zero c in K. This implies $(f^-)^\sigma = cf^-$, and taking the product shows that $c^2 = 1$. Writing $(A + YB)^2 = A_1 + YB_1$ for $A_1 = A^2 + DB^2$, $B_1 = 2AB$ we deduce that $A_1, B_1 \neq 0$ are in $\mathbb{C}(t)[X]$ also with $A_1^2 - DB_1^2 = 1$. Clearing denominators we find $A_2, B_2 \neq 0$ in $\mathbb{C}[t, X]$ and c_2 in $\mathbb{C}[t]$ with $A_2^2 - (X^6 + X + t)B_2^2 = c_2$. If c_2 is in $\mathbb{C}$ we get an immediate contradiction on examining the coefficients of the highest power of t. Otherwise specializing t to a zero t_0 of c_2 would show that $X^6 + X + t_0$ is a square in $\mathbb{C}(X)$, clearly impossible.

It follows from Lemma 10.1(iii) that the point $\{\beta(\infty^+), \beta(\infty^+)\}$ is not identically torsion. We are all set up to apply our Theorem, or more conveniently the Proposition directly, with C a suitable curve corresponding to the splitting field of $X^6 + X + t$ (in fact a 720-fold cover of $\mathbb{A}$ deprived of the 5 points with $46656t^5 = 3125$). But first we should know that the Jacobian is generically simple, and it suffices to show that the Jacobian J_D of $Y^2 = X^6 + X + t$ is generically simple. We will do this by showing that there are infinitely many t_0 such that the Jacobian of $Y^2 = X^6 + X + t_0$ is simple.

We use the criterion of Stoll [St] explained (with a misprint) in [CF, p. 158]. The curve $Y^2 = X^6 + X + 1$ has discriminant -43531 not divisible by 7. It has 9 points over the finite field $\mathbb{F}_7$ (including ∞^+, ∞^-). Similarly it has 67 points over $\mathbb{F}_{49}$. In the notation of [CF] we calculate $a_7 = -1$, $b_7 = 9$ leading to the test polynomial

$$C(T) = T^4 + \tfrac{5}{7}T^3 - \tfrac{3}{7}T^2 + \tfrac{5}{7}T + 1.$$

So $a_7^2 - 4(b_7 - 14) = 21$ is not a square in $\mathbb{Q}$. And it is easily checked that $C(\zeta) \neq 0$ for any root of unity ζ with $\zeta^n = 1$ for $n = 1, 2, 3, 4, 5, 6, 8, 10, 12$ (for example the resultants with $T^n - 1$ are non-zero). Thus the Jacobian of $Y^2 = X^6 + X + 1$ is simple. But the same calculation shows that this also holds for any $Y^2 = X^6 + X + t_0$ with t_0 congruent to 1 modulo 7. This suffices for the identical simplicity (where incidentally [ST] is implicitly used in the form of the isogeny-invariance of good reduction, as in Corollary 2 (p. 493) there, to see that the possible elliptic curves whose product is isogenous to the Jacobian both have good reduction themselves).

This completes the proof of Theorem P1 that there are at most finitely many complex values of t for which the Pell equation for $D = X^6 + X + t$ is solvable. As pointed out in (1.6), this holds for $t = 0$. We mentioned that it does not hold for $t = 1$; here is a proof.

When k is a finite field, the continued fraction method over $k[X]$ does work just as for $\mathbb{Z}$ (see for example [PT, p. 157]); the expansions of square roots are always periodic and the Pell equation (1.5) is always solvable. Olaf Merkert has calculated the minimal solutions for $k = \mathbb{F}_3$ and $k = \mathbb{F}_5$ with $D = X^6 + X + 1$. For $\mathbb{F}_3$ he finds

$$A = 2X^{14} + X^{12} + X^{10} + X^9 + X^8 + X^7 + 2X^6 + 2X^5 + 2X^4 + X^3 + X^2 + 2$$

of degree 14 and

$$B = 2X^{11} + X^9 + X^7 + 2X^4 + X.$$

For $\mathbb{F}_5$ he finds

$$A = 2X^{31} + X^{30} + 3X^{29} + X^{28} + X^{25} + 2X^{24} + 3X^{22} + 3X^{21} + 3X^{20} + X^{19} + 4X^{17} + X^{16} + 4X^{15} + 4X^{13} + 2X^{12} + 2X^{11} + X^{10} + X^8 + 3X^7 + X^6 + 3X^5 + 2X^4 + 3X^3 + 2X^2 + 4$$

of degree 31 and

$$B = 2X^{28} + X^{27} + 3X^{26} + X^{25} + 4X^{23} + 2X^{22} + 3X^{20} + 4X^{16} + 2X^{15} + 4X^{13} + 2X^{12} + 2X^{11} + X^{10} + X^8 + 3X^7 + 3X^6 + 4X^5 + 2X^3 + X^2 + 4X.$$

Now suppose (1.5) is solvable over $\mathbb{C}[X]$ for $D = X^6 + X + 1$. Then A is not constant and so by Lemma 10.1(i) with $K = \mathbb{C}$ the point Π_0 on the Jacobian J_D corresponding to the divisor $(\infty^+) - (\infty^-)$ is torsion; let n_0 be its order.

Write $n_0 = 3^e m$ for a non-negative integer e and an integer m prime to 3, and consider the point $\Lambda_0 = 3^e 28\Pi_0$ on J_D. As the discriminant -43531 above is not divisible by 3, we can reduce modulo 3 to get $\tilde{\Lambda}_0 = 3^e(28\tilde{\Pi}_0)$ on the abelian variety $\tilde{J}_D$. However, by Merkert's calculation and Lemma 10.1(i) with $K = \mathbb{F}_3$ we see that $28\tilde{\Pi}_0 = 0$. Thus $\tilde{\Lambda}_0 = 0$. But Λ_0 has order dividing m which is prime to 3; hence again by [ST] we deduce that $\Lambda_0 = 0$ (see for example [ST, Lemma 2, p. 495] and the short paragraph immediately following the proof). Therefore the order n_0 divides $3^e 28$.

A similar argument over $\mathbb{F}_5$ shows that n_0 divides $5^f 62$ for some non-negative integer f.

It follows that n_0 divides 2, so that $2\Pi_0 = 0$. But then there is a function $A(X) + YB(X)$ on H_D with a zero of order 2 at ∞^+ and a pole of order 2 at ∞^-, and no other

zeros or poles. Now Lemma 10.1(ii) with $K = \mathbb{C}$ yields A, B in $\mathbb{C}[X]$ with A of degree 2 such that $A^2 - DB^2 = c$ for some $c \neq 0$ in $\mathbb{C}$, a clear impossibility.

We get the same conclusion for the t for which the continued fraction of

$$\sqrt{X^6 + X + t} = X^3 \sum_{k=0}^{\infty} \binom{1/2}{k} (X^{-5} + tX^{-6})^k = X^3 + \frac{1}{2}X^{-2} + \frac{t}{2}X^{-3} - \frac{1}{8}X^{-7} + \cdots \tag{10.3}$$

is periodic. In the usual notation $[a_0; a_1, a_2, \ldots]$ it starts

$$a_0 = X^3, \quad a_1 = 2X^2 - 2tX + 2t^2, \quad a_2 = -\frac{1}{2t^3}X - \frac{1}{2t^2}, \quad a_3 = -8t^6X + 16t^7 \quad (t \neq 0) \tag{10.4}$$

(so it is not "continuous in t"). And we get the same conclusion for the t for which there exists a non-zero complex polynomial E of degree at most 4 such that $E(X)/\sqrt{X^6 + X + t}$ is integrable in elementary terms.

It is rather likely that similar arguments could be carried out for $D = D(X) = F(X)(X - t)$ with fixed quintic F defined over the field of algebraic numbers. The family $Y^2 = D(X)$ is isomorphic to the family $y^2 = f(x)(x - s)$ via

$$x = \frac{1}{X - \alpha}, \qquad y = \frac{Y}{(X - \alpha)^3}, \qquad s = \frac{1}{t - \alpha}$$

for any zero α of F, with quartic $f(x) = x^5 F(1/x + \alpha)$. In [EEHK] Ellenberg, Elsholtz, Hall and Kowalski show for example that the Jacobian of the second family is identically simple (and even that the endomorphism ring is $\mathbb{Z}$).

But what goes wrong for $D = X^6 + X^2 + t$? The argument above still shows that Pell's equation is not solvable identically. However, by [CF, Theorem 14.1.1(i), p. 155] we see that the Jacobian is not identically simple. In fact there are maps β_1, β_2 defined by

$$\beta_1(X, Y) = (X_1, Y_1) = (X^2, Y), \qquad \beta_2(X, Y) = (X_2, Y_2) = (X^2, XY)$$

from H_D to elliptic curves E_1, E_2 defined respectively by

$$Y_1^2 = X_1^3 + X_1 + t, \qquad Y_2^2 = D_2(X_2) = X_2^4 + X_2^2 + tX_2.$$

We have $\beta_1(\infty^{\pm}) = \infty_1$, $\beta_2(\infty^{\pm}) = \infty_2^{\pm}$ for the point at infinity on E_1 and the two points at infinity on E_2. Thus $(\infty^+) - (\infty^-)$ projects down to something identically torsion on E_1, and to $(\infty_2^+) - (\infty_2^-)$ on E_2. This enables us to use the arguments of Lemma 10.1 for genus 1 instead of 2. In fact if some $A_2(X_2) - Y_2 B_2(X_2)$ has suitable zeros and poles at $\infty_2^{\pm}$ on E_2 then we can pull it back to get $A_2(X^2) - XYB_2(X^2)$ with suitable zeros and poles at $\infty^{\pm}$; and indeed from $1 = A_2(X_2)^2 - D_2(X_2)B_2(X_2)^2$ we get

$$\begin{aligned} 1 &= A_2(X^2)^2 - D_2(X^2)B_2(X^2)^2 = A^2 - (X^8 + X^4 + tX^2)B_2(X^2)^2 \\ &= A^2 - (X^6 + X^2 + t)B^2 \end{aligned}$$

for $A = A_2(X^2)$, $B = XB_2(X^2)$.

Incidentally it may be shown that the map ι from $J_{\lambda\kappa\theta}$ to $E_1 \times E_2$ defined by

$$\iota(\{P, Q\}) = \big(\beta_1(\beta^{-1}(P)) + \beta_1(\beta^{-1}(Q)),\ \beta_2(\beta^{-1}(P)) + \beta_2(\beta^{-1}(Q))\big)$$

is an isogeny (compare [CF, p. 155]). A simple calculation shows it is of degree 4. And the curves E_1, E_2 are not isogenous; for example their j-invariants are

$$j_1 = \frac{6912}{27t^2+4}, \qquad j_2 = -\frac{256}{t^2(27t^2+4)},$$

so j_2 cannot be integral over $\mathbb{C}[j_1]$ (as would be predicted by the classical theory) because of its pole at $t = 0$.

It is not difficult to see that there are infinitely many complex values of t for which $(\infty_2^+) - (\infty_2^-)$ is torsion on E_2; several methods are discussed in [Za, Notes to Chapter 3, p. 92]. So there are also infinitely many complex values of t for which the Pell equation

$$A^2 - (X^6 + X^2 + t)B^2 = 1$$

is solvable. For example we can reduce to Weierstrass form with the map ζ defined by

$$\zeta(X_2, Y_2) = (W, Z) = \left(\frac{t}{X_2} + \frac{1}{3},\ \frac{tY_2}{X_2^2}\right)$$

from E_2 to the curve E defined by $Z^2 = W^3 + uW + v$ for $u = -1/3,\ v = t^2 + 2/27$. We have $\zeta(\infty_2^\pm) = P^\pm = (1/3, \pm t)$. In terms of the denominator $B_n(W, u, v)$ of the classical rational function describing multiplication on E by n, the t are precisely the zeros of $B_n(1/3, -1/3, t^2 + 2/27)$ $(n = 1, 2, \ldots)$

A reasonable explicit value is $t = i/2$ with $n = 5$. This comes from the function $g^+ = a + bW + cW^2 + Z(d + eW)$ with a zero of order 5 at P^+ and a pole of order 5 at infinity, where

$$a = 10i, \qquad b = 3i, \qquad c = -18i, \qquad d = -24, \qquad e = 18.$$

Thus $g^- = a + bW + cW^2 - Z(d + eW)$ has a zero of order 5 at P^- and a pole of order 5 at infinity. So g^+/g^- has a zero of order 5 at P^+ and a pole of order 5 at P^-. Pulling this back to E_2 and then to H_D, we end up with

$$A = 16iX^{10} + 16X^8 + 12iX^6 + 8X^4 + 4iX^2 + 1, \qquad B = 16iX^7 + 16X^5 + 4iX^3 + 4X$$

satisfying

$$A^2 - (X^6 + X^2 + i/2)B^2 = 1,$$

with A of degree 10 as predicted by Lemma 10.1. Or in terms of continued fractions, with the usual notation for the period,

$$\sqrt{X^6 + X^2 + i/2} = [X^3; \overline{2X, iX, -4iX^3, iX, 2X, 2X^3}].$$

Truncating before the final $2X^3$ gives A, B as above. But truncating before the $-4iX^3$ leads to the smaller solution

$$A_0 = (2-2i)X^5 - (1+i)X^3 + (1-i)X, \qquad B_0 = (2-2i)X^2 - (1+i)$$

of degree 5, and the solution above is up to sign the "square".

And indeed

$$\int \frac{-(10X^2+2i)}{\sqrt{X^6+X^2+\frac{i}{2}}}\,dX = \log\left(A+B\sqrt{X^6+X^2+\frac{i}{2}}\right) = 2\log\left(A_0+B_0\sqrt{X^6+X^2+\frac{i}{2}}\right)$$

(up to constants) as predicted by (10.2).

11. Almost the Pell equation

When solving a Pell equation $a^2 - db^2 = 1$ over $\mathbb{Z}$ one notes that a/b must be a good rational approximation to $\sqrt{d}$. But constructing such good approximations by the Box Principle gives infinitely many solutions only of the equation $a^2 - db^2 = m$ for some fixed m, "almost the Pell equation". To obtain $m = 1$ an extra application of the Box Principle is needed.

Analogous considerations for $A^2 - DB^2 = 1$ over $K[X]$ lead also to an equation $A^2 - DB^2 = M$; but here M is not fixed, merely of degree at most 2. See for example [PT, p. 157]. Now there is no general way to obtain $M = 1$, and indeed we have seen that this is impossible for $D = X^6 + X + t$ for all but finitely many complex values of t.

Indeed in this parametric situation the resulting M, listed somehow as M_n ($n = 1, 2, \ldots$), can be assumed to have the form $c_n''(t)X^2 + c_n'(t)X + c_n(t)$ for $c_n(t)$, $c_n'(t)$, $c_n''(t)$ in $K(t)$. Then we would have to solve the equations $c_n''(t) = c_n'(t) = 0$ for t. This is another illustration of the term "unlikely intersection", such as in the very simplest example $t^n = (1-t)^n = 1$ in $\mathbb{G}_{\mathrm{m}}^2$, and does indeed lead to a solution set that is at most finite.

However, the equations $c_n''(t) = 0$ alone, just like $t^n = 1$, are not unlikely in this sense; and one would expect them to have infinitely many solutions as n varies. For our special D above this leads to

$$A^2 - (X^6 + X + t)B^2 = c'X + c. \tag{11.1}$$

For us this is "almost the Pell equation" over $\mathbb{C}[X]$, as in Theorem P2.

A simple example is $t = 0$ with $A = X^3$, $B = 1$ and

$$A^2 - (X^6 + X)B^2 = -X. \tag{11.2}$$

Less simple is $t = \sqrt[5]{1/12}$ with

$$A = 24t^3X^7 - 48t^4X^6 + 6X^5 - 6tX^4 + 6t^2X^3 + 12t^3X^2 - 12t^4X + 1,$$
$$B = 24t^3X^4 - 48t^4X^3 + 6X^2 - 6tX + 6t^2,$$

and

$$A^2 - (X^6 + X + t)B^2 = 12t^4X - 2.$$

It may be shown that such values of t are precisely those that occur as poles of the partial quotients in $a_1, a_2, \ldots$ in (10.4). We see at once that $t = 0$ in a_2, and also $t = \sqrt[5]{1/12}$ by going further to

$$a_4 = -\frac{X}{2t^3(12t^5 - 1)} + \frac{16t^5 - 1}{4t^2(12t^5 - 1)}.$$

Similar issues occurred in the problems considered in [MZ1], [MZ2]; the pencil of abelian surfaces of this paper was there replaced by the square of the Legendre family. There the unlikely intersection corresponded to the equations

$$n\big(2, \sqrt{2(2-\lambda)}\big) = n\big(3, \sqrt{6(3-\lambda)}\big) = 0$$

on the Legendre elliptic curve as in (1.1). And there we also considered a "likely intersection" with $n(2, \sqrt{2(2-\lambda)}) = 0$ alone. We proved that there are infinitely many λ using a very special case of Siegel's theorem on integral points on curves over function fields. Other proofs were later presented (see [Za, pp. 92, 93]), but the matter, although not difficult, seemed not completely obvious.

In this case of a simple abelian surface family, things appear to be more complicated. The approach through Siegel's theorem seems to require a deeper analogue for integral points in affine subsets of abelian variety. (This is due to Faltings in the number field case. Here we would need the function field analogue; presumably, although less deep, this should be still much more difficult compared to Siegel's theorem.) On the other hand, the elliptic case admitted also an analytical approach (working on complex tori rather than on algebraic models), which was simple and moreover gave additional information. It is an approach of this nature that we shall adopt here.

We now start the proof of Theorem P2. More precisely, we shall prove that given any d_0, there are infinitely many complex t for which there exist complex $c' \neq 0$, c and A, B in $\mathbb{C}[X]$ with A of degree at least d_0 and (11.1).

We start with an analogue of Lemma 10.1 for (11.1), as there over $K[X]$, where now D is squarefree of degree 6. We embed $H_{\lambda\kappa\theta}$ in $J_{\lambda\kappa\theta}$ by mapping R to

$$j(R) = \{\beta(\infty^+), \beta(\infty^+)\} - \{\beta(\infty^+), R\}. \tag{11.3}$$

Write $V_{\lambda\kappa\theta}$ for the image $j(H_{\lambda\kappa\theta})$; it is an algebraic curve but we do not need to know this. It contains the origin $j(\beta(\infty^+))$; and after removing this we write temporarily $\hat{V}_{\lambda\kappa\theta}$ for what is left.

Lemma 11.1. (i) *Suppose there are A, B in $K[X]$ and $c' \neq 0, c$ in K with $A \neq 0$ of degree d such that $A^2 - DB^2 = c'X + c$. Then for $n = d$ or $1 - d$ the function $A(X) + YB(X)$ on H_D has a zero of order $n - 1$ at ∞^+, a pole of order n at ∞^-, and one other zero $\gamma^+ \neq \infty^+$ at which $c'X + c$ also vanishes.*

(ii) *Suppose there are A, B in $K(X)$ and n such that $A(X) + YB(X)$ on H_D has a zero of order $n-1$ at ∞^+, a pole of order n at ∞^-, and one other zero $\gamma^+ \neq \infty^+$. Then $n = d$ or $1-d$ for some integer $d \geq 0$ and A, B are in $K[X]$ with $A \neq 0$ of degree d such that $A^2 - DB^2 = c'X + c$ for some $c' \neq 0$, c in K such that $c'X + c$ also vanishes at γ^+.*

(iii) *Suppose D splits completely over K. There are A, B in $K[X]$ and $c' \neq 0$, c in K with $A \neq 0$ of degree d such that $A^2 - DB^2 = c'X + c$ if and only if $d \geq 1$ and the point $d\{\beta(\infty^+), \beta(\infty^+)\}$ is in $\hat{V}_{\lambda\kappa\theta}$.*

Proof. In the situation of (i) we have $f^+ f^- = c'X + c$ for the functions $f^+ = A(X) + YB(X)$, $f^- = A(X) - YB(X)$ on H_D. So the only possible zeros and poles are at ∞^+, ∞^- and the two zeros γ^+, γ^- (possibly coinciding) of $c'X + c$. Since the number of zeros is the number of poles, there is an integer n such that f^+ has a zero of order $n-1$ at ∞^+, a simple zero at γ^+, and a pole of order n at ∞^-. Now $f^+ + f^- = 2A$ has poles of order n or $1-n$ at ∞^+, ∞^- and no other poles. Thus this order is d, and that proves (i).

In the situation of (ii) with $f^+ = A(X) + YB(X)$ we deduce that $f^- = A(X) - YB(X)$ has a pole of order n at ∞^+, a zero of order $n-1$ at ∞^-, and a simple zero at γ^-. Thus $f^+ f^- = c'X + c$ for constants $c' \neq 0$, c. Also $f^+ + f^- = 2A$ has no poles at finite points, so it must be a polynomial, clearly non-zero because $c' \neq 0$. And finally because D is squarefree we see from $DB^2 = A^2 - c'X - c$ that B is also a polynomial. This brings us back to (i) and thereby completes the proof of (ii).

We may note that in the above situations the points $\gamma^\pm = (X_0, \mp A(X_0)/B(X_0))$ for $X_0 = -c/c'$ are defined over K.

Finally in the situation of (iii) the existence of A, B clearly implies $d \geq 1$ and gives from (i) a function $f^+ = A(X) + YB(X)$ from which we derive a function g on $H_{\lambda\kappa\theta}$ with a zero of order $n-1$ at $P^+ = \beta(\infty^+)$, a pole of order n at $P^- = \beta(\infty^-)$, and a simple zero at $Q^+ = \beta(\gamma^+) \neq P^+$. Here $n = d$ or $1-d$; but by changing the sign of B we can assume $n = d$. Looking at gl^n as in the proof of Lemma 10.1 now shows that $n\{P^+, P^+\} = j(Q^+)$. And conversely if $n\{P^+, P^+\} = j(Q)$ for some $n = d \geq 1$ and some $Q \neq P^+$, then going backwards we find a function on H_D with a zero of order $n-1$ at ∞^+, a simple zero at $\gamma^+ = \beta^{-1}(Q) \neq \infty^+$, and a pole of order n at ∞^-. This can be written as $f^+ = A(X) + YB(X)$ for A, B in $K(X)$, and we are back to (ii). That completes the proof of the present lemma.

Now for the proof concerning (11.1) we see that we are in a situation like that of the Proposition, except that the condition of the point $\{P(\mathbf{c}), Q(\mathbf{c})\}$ being torsion on $J_{\lambda(\mathbf{c})\kappa(\mathbf{c})\theta(\mathbf{c})}$, that is, $n\{P(\mathbf{c}), Q(\mathbf{c})\} = O$, is replaced by $n\{P(\mathbf{c}), Q(\mathbf{c})\}$ lying on the curve $\hat{V}_{\lambda(\mathbf{c})\kappa(\mathbf{c})\theta(\mathbf{c})}$ (for $n \geq 1$). Here of course $P(\mathbf{c}) = Q(\mathbf{c}) = \beta(\infty^+)$. The latter curve being of positive dimension, the corresponding condition is much less stringent and we will prove that for each n_0 it holds for infinitely many $\mathbf{c}$ in C with some $n \geq n_0$. In fact it would suffice to land in the hatless $V_{\lambda(\mathbf{c})\kappa(\mathbf{c})\theta(\mathbf{c})}$ (subsequently written $V(\mathbf{c})$ for brevity) in $J_{\lambda(\mathbf{c})\kappa(\mathbf{c})\theta(\mathbf{c})}$ (subsequently written $J(\mathbf{c})$ for brevity) because of Theorem P1; however, we prefer not to evoke this rather deeper result here. Incidentally, we do not need to assume anything for the generic point of C as we did in the Proposition, because some $n \geq 1$ with

$n\{P, Q\}$ in $V_{\lambda\kappa\theta}$ identically now works in our favour: by specialization it gives infinitely many $\mathbf{c}$ with the same n. In fact the identity (1.9), which for convenience we display again as

$$A^2 - (X^6 + X + t)B^2 = -X - t \tag{11.4}$$

with $A = X^3,\ B = 1$, is an example of this with $n = 3$. We will verify later on that this is essentially the only generic example; thus one cannot obtain $\mathbf{c}$ with $n \geq 4$ simply by this sort of specialization (e.g. to $t = 0$ as in (11.2) above).

Our general strategy may be sketched as follows. We work near some suitable fixed point $\mathbf{c}_*$ on C. First we find a large n with $n\{P(\mathbf{c}_*), Q(\mathbf{c}_*)\}$ near zero in $V(\mathbf{c}_*)$. Then we perturb $\mathbf{c}_*$ by an amount of order $1/n$, staying on C, to some $\mathbf{c}_{**}$. As "$nf(x + y/n)$ is about $nf(x) + yf'(x)$" we can by suitable choice of $\mathbf{c}_{**}$ bring $n\{P(\mathbf{c}_{**}), Q(\mathbf{c}_{**})\}$ near a better behaved point of $V(\mathbf{c}_*)$ and so near a better behaved point of $V(\mathbf{c}_{**})$. Then another perturbation to $\mathbf{c}$ places $n\{P(\mathbf{c}), Q(\mathbf{c})\}$ exactly on $V(\mathbf{c})$ (but non-zero) as required in Lemma 11.1(iii); this last step involves some form of implicit function theorem which requires the better behaviour.

In fact we linearize the procedure using tangent spaces. Thus we will need the period functions $\mathbf{f}, \mathbf{g}, \mathbf{k}, \mathbf{l}$ of Section 4 together with an abelian logarithm $\mathbf{z}$ of $\{P, Q\}$ as in (4.3). These were defined first at all $\mathbf{c}$ near some $\mathbf{c}_*$ as in (4.2). For each $\mathbf{c}$ these periods generate the lattice $\Omega = \Omega(\mathbf{c})$ over $\mathbb{Z}$. We will work with the inverse image $Z(\mathbf{c})$ of $V(\mathbf{c})$ under the exponential map; by (11.3) and (4.3) this consists of

$$z_R(\mathbf{c}) = \left(\int_R^{P(\mathbf{c})} \frac{dX}{Y},\ \int_R^{P(\mathbf{c})} \frac{X\,dX}{Y} \right)$$

taken over all possible R and all possible paths. By Riemann's Theorem it is the zero-set of a suitable theta-function ϑ. In fact we have

$$\vartheta\left(\tfrac{1}{2}z_\infty(\mathbf{c}) - \tfrac{1}{2}z_R(\mathbf{c}); \mathcal{T}\right) = 0 \tag{11.5}$$

for some matrix $\mathcal{T} = \mathcal{T}(\mathbf{c})$ in the Siegel upper half-space (see for example [G, p. 98]). Thus $Z(\mathbf{c})$ is a complex analytic curve in $\mathbb{C}^2$ containing $\Omega(\mathbf{c})$, and it is everywhere smooth, being locally analytically isomorphic to $V(\mathbf{c})$ and so to $H(\mathbf{c}) = H_{\lambda(\mathbf{c})\kappa(\mathbf{c})\theta(\mathbf{c})}$ (or by Riemann's Singularity Theorem for genus 2). It is known to be connected.

For the moment we will work with just $J(\mathbf{c}_*)$, $V(\mathbf{c}_*)$, $\Omega(\mathbf{c}_*)$, $Z(\mathbf{c}_*)$, which for further brevity we will denote by J_*, V_*, Ω_*, Z_* respectively. By adjusting $\mathbf{c}_*$ we can assume that J_* is simple (for example we could use the t_0 obtained from Stoll's criterion in Section 10, or a general result [M2] of the first author). Later on we will make another adjustment of this type.

Lemma 11.2. *Given $\mathbf{u} \neq 0$ in $\mathbb{C}^2$ there is $\mathbf{z}_*$ in $Z_* \setminus \Omega_*$, and also in the topological closure of $\mathbb{C}\mathbf{u} + \Omega_*$, such that the tangent space of Z_* at $\mathbf{z}_*$ does not contain $\mathbf{z}_* + \mathbf{u}$.*

Proof. The closure U of $\mathbb{C}\mathbf{u} + \Omega_*$ in $\mathbb{C}^2$ (i.e. $\mathbb{R}^4$), as for any group in a real vector space, must have the form $G + S$ for a group G and a real vector subspace S, of dimension say s, with G discrete in $\mathbb{R}^4/S$. As U contains Ω_* we see that S contains a subgroup of Ω_* of

rank s. As U contains $\mathbb{C}\mathbf{u}$, we must have $s \geq 2$. But $s = 2$ would give an elliptic curve in J_*, contradicting its simplicity.

First we want to show that $Z_* \cap S$ modulo Ω_* is infinite. This is clear if $s = 4$, so we assume $s = 3$.

Now the removal of S (i.e. $\mathbb{R}^3$) disconnects $\mathbb{C}^2$ (i.e. $\mathbb{R}^4$). As Z_* is connected and contains Ω_*, it follows that $Z_* \cap S$ is not empty. If this were finite modulo Ω_* then it would be a discrete set of points in $\mathbb{C}^2$, and removing these from Z_* would still leave a connected set $\hat{Z}_*$. This still contains some translate of Ω_*, and the argument above would give something in the empty set $\hat{Z}_* \cap S$. Thus indeed $Z_* \cap S$ modulo Ω_* is infinite; and we can find an infinite subset T of $Z_* \cap S$ lying in a compact subset of $\mathbb{C}^2$.

Now if the complex tangent line of Z_* through $\mathbf{t}$ contains $\mathbf{t} + \mathbf{u}$ for all $\mathbf{t}$ in $T \setminus \Omega_*$ then this would be the case identically in $\mathbf{z}$ on the complex analytic curve Z_*. That would imply that Z_* is a complex line. But then it cannot contain Ω_* (for example by the simplicity of J_*). This completes the proof.

We will choose $\mathbf{u}$ (and so $\mathbf{z}_*$ in Z_*) later on; they will depend only on the choice of $\mathbf{c}_*$. We choose in a similar way also a small $\varepsilon > 0$, say with $\varepsilon \leq 1$. The lemma implies that there is a period $\mathbf{w}_*$ in Ω_* and τ in $\mathbb{C}$ with

$$|\mathbf{z}_* - \tau\mathbf{u} - \mathbf{w}_*| < \varepsilon. \tag{11.6}$$

As 0 is a cluster point of $\mathbb{N}\mathbf{u}$ modulo Ω_*, we can adjust τ by an integer so that $|\tau| \geq 1$.

Now $\mathbb{N}\mathbf{z}(\mathbf{c}_*)$ also clusters near 0 modulo Ω_*, and so there are infinitely many natural numbers n for which there exists a period $\mathbf{w}^\sharp_{*n}$ in Ω_* with

$$|n\mathbf{z}(\mathbf{c}_*) - \mathbf{w}^\sharp_{*n}| < \varepsilon. \tag{11.7}$$

We can assume that $n \geq n_0$ for our prescribed n_0, and we can also assume that

$$n \geq (1 + |\tau|^2)/\varepsilon. \tag{11.8}$$

We are now going to move $\mathbf{c}$ on C slightly away from $\mathbf{c}_*$; we do this by choosing any non-constant function on C—the coefficient in $X^6 + X + t$ will do perfectly—and regarding $\mathbf{c} = \mathbf{c}(t)$ as a function of t near $t_* = t(\mathbf{c}_*)$. Thus we shall write $\tilde{\mathbf{z}}(t) = \mathbf{z}(\mathbf{c}(t))$ etc. The point is that a first approximation to $n\tilde{\mathbf{z}}(t_* + \tau/n)$ is $n\tilde{\mathbf{z}}(t_*) + \tau\tilde{\mathbf{z}}'(t_*)$, where the prime denotes d/dt; thus although the perturbation τ/n on t may be small, the effect on $n\tilde{\mathbf{z}}$ may not be. We have in fact

$$n\tilde{\mathbf{z}}(t_* + \tau/n) = n\tilde{\mathbf{z}}(t_*) + \tau\tilde{\mathbf{z}}'(t_*) + O(|\tau|^2/n) \tag{11.9}$$

where the implicit constant, as all such constants below, is independent of τ, ε and especially n. In fact it will be seen that they can be taken as absolute constants, provided t_* is chosen in a fixed way (and we can almost certainly take $t_* = 2$, for example).

This enhanced perturbation will enable us to deduce from (11.7) that $n\tilde{\mathbf{z}}(t_* + \tau/n)$ is close not to zero but to $\mathbf{z}_*$ in Z_*, at least up to periods. However, we must take into account the effect of perturbing the lattice. Writing

$$\mathbf{w}^\sharp_{*n} = p_n\tilde{\mathbf{f}}(t_*) + q_n\tilde{\mathbf{g}}(t_*) + r_n\tilde{\mathbf{k}}(t_*) + s_n\tilde{\mathbf{l}}(t_*)$$

for integers p_n, q_n, r_n, s_n we therefore define

$$\mathbf{w}_n^\sharp = p_n\tilde{\mathbf{f}}(t_* + \tau/n) + q_n\tilde{\mathbf{g}}(t_* + \tau/n) + r_n\tilde{\mathbf{k}}(t_* + \tau/n) + s_n\tilde{\mathbf{l}}(t_* + \tau/n)$$

in $\tilde{\Omega}(t_*+\tau/n)$. To estimate $\mathbf{w}_n^\sharp - \mathbf{w}_{*n}^\sharp$ we have to be careful about the sizes of p_n, q_n, r_n, s_n. In fact we can write

$$\tilde{\mathbf{z}}(t_*) = x\tilde{\mathbf{f}}(t_*) + y\tilde{\mathbf{g}}(t_*) + u\tilde{\mathbf{k}}(t_*) + v\tilde{\mathbf{l}}(t_*) \tag{11.10}$$

for real x, y, u, v, and from (11.7) it follows (since period lattices are discrete) that

$$p_n = nx + O(\varepsilon),\ q_n = ny + O(\varepsilon),\ r_n = nu + O(\varepsilon),\ s_n = nv + O(\varepsilon). \tag{11.11}$$

We find that

$$\mathbf{w}_n^\sharp = \mathbf{w}_{*n}^\sharp + \tau\mathbf{h} + O(|\tau|^2/n) + O(|\tau|\varepsilon/n) \tag{11.12}$$

where

$$\mathbf{h} = x\tilde{\mathbf{f}}'(t_*) + y\tilde{\mathbf{g}}'(t_*) + u\tilde{\mathbf{k}}'(t_*) + v\tilde{\mathbf{l}}'(t_*) \tag{11.13}$$

may possibly be related to quasi-periods. An analogous construction produces $\mathbf{w}_n$ in $\tilde{\Omega}(t_*+\tau/n)$ from $\mathbf{w}_*$ in (11.6) but now with coefficients $O(|\tau|)$ instead of those in (11.11), and we find

$$\mathbf{w}_n = \mathbf{w}_* + O(|\tau|^2/n). \tag{11.14}$$

We now choose

$$\mathbf{u} = \tilde{\mathbf{z}}'(t_*) - \mathbf{h}. \tag{11.15}$$

By bad luck it may happen that $\mathbf{u} = 0$, but if so we can just modify the choice of t_* to get $\mathbf{u} \neq 0$. We postpone the details of this step until later; they rely on our algebraic independence result (Lemma 5.1). Anyway, when we combine (11.7) with (11.6), (11.9), (11.12), (11.14) we see that $n\tilde{\mathbf{z}}(t_* + \tau/n)$ is close to $\mathbf{z}_*$ modulo $\tilde{\Omega}(t_* + \tau/n)$; and using (11.8) to tidy up the error terms we end up with

$$n\tilde{\mathbf{z}}(t_* + \tau/n) = \mathbf{z}_* + \mathbf{w}_n^\sharp - \mathbf{w}_n + O(\varepsilon). \tag{11.16}$$

Here $\mathbf{z}_*$ was on $Z_* = \tilde{Z}(t_*)$. The last step is to make an additional perturbation from $t_* + \tau/n$ to

$$t_n = t_* + \tau/n + \zeta_n/n \tag{11.17}$$

so as to have $n\tilde{\mathbf{z}}(t_n)$ actually on $\tilde{Z}(t_n)$, at first modulo periods but then since the $\tilde{Z}(t)$ are periodic this is good enough. Here $|\zeta_n| \leq 1$ to start with.

For this we use the fact that near $(\mathbf{z}_*, t_*)$ the set of $(\mathbf{z}, t)$ in $\mathbb{C}^2 \times \mathbb{C}$ with $\mathbf{z}$ in $\tilde{Z}(t)$ is defined locally by an analytic equation $f(\mathbf{z}, t) = 0$. This seems to be well-known; for example the $\mathcal{T} = \mathcal{T}(\mathbf{c})$ in (11.5) is analytic in $\mathbf{c}$ (see also [G, p. 97]) and so in t. Again we must adjust the lattice, and so first we define the periods

$$\mathbf{w}_n^\sharp(\zeta) = p_n\tilde{\mathbf{f}}(t_n) + q_n\tilde{\mathbf{g}}(t_n) + r_n\tilde{\mathbf{k}}(t_n) + s_n\tilde{\mathbf{l}}(t_n)$$

and analogously $\mathbf{w}_n(\zeta)$ in $\tilde{\Omega}(t_n)$. Then we define

$$F_n(\zeta) = f(n\tilde{\mathbf{z}}(t_n) - \mathbf{w}_n^\sharp(\zeta) + \mathbf{w}_n(\zeta), t_n).$$

Note that the estimates (11.9), (11.12), (11.14) with $\tau + \zeta$ in place of τ, together with (11.16), show that the first expression in f is

$$n\tilde{\mathbf{z}}(t_n) - \mathbf{w}_n^\sharp(\zeta) + \mathbf{w}_n(\zeta) = \mathbf{z}_* + O(|\zeta|) + O(\varepsilon); \tag{11.18}$$

and the second expression is even $t_* + O(\varepsilon)$. So F_n is well-defined provided ζ, ε are sufficiently small.

We wish to find ζ_n with $F_n(\zeta_n) = 0$. We apply the Rouché theorem (see for example [L, p. 158]) to the functions $F_n(\zeta)$ and $F_n(\zeta) - F_n(0)$; if we can verify that $|F_n(0)| < |F_n(\zeta) - F_n(0)|$ on $|\zeta| = \rho$ for a suitable radius $\rho \le 1$, then because the second function has the zero $\zeta = 0$, we get the required zero of the first function.

To start with, by (11.16) we have

$$F_n(0) = O(\varepsilon). \tag{11.19}$$

Next with g_n being the first derivative of $F_n(\zeta)$ at $\zeta = 0$ we can verify that

$$F_n(\zeta) - F_n(0) = \zeta g_n + O(|\zeta|^2) \tag{11.20}$$

when for safety $|\zeta| \le 1/2$—for example by using the Cauchy integral formula for $\frac{\zeta^2}{2\pi i}\int \frac{F_n(z)\,dz}{z^2(z-\zeta)}$ over $|z| = 3/4$ and estimating $F_n(z) = O(1)$. Here g_n is given by

$$\mathbf{d}f(n\tilde{\mathbf{z}}(t_*+\tau/n)-\mathbf{w}_n^\sharp+\mathbf{w}_n, t_*+\tau/n)\cdot(\tilde{\mathbf{z}}'(t_*+\tau/n)-(\mathbf{w}_n^\sharp)'(0)+\mathbf{w}_n'(0))+\tfrac{1}{n}f_t(t_*+\tau/n),$$

where $\mathbf{d}f$ is the gradient with respect to $\mathbf{v}$, f_t the derivative with respect to t, and the dot is the scalar product. By (11.16) and (11.8) we have

$$\mathbf{d}f(n\tilde{\mathbf{z}}(t_* + \tau/n) - \mathbf{w}_n^\sharp + \mathbf{w}_n, t_* + \tau/n) = \mathbf{d}f(\mathbf{z}_*, t_*) + O(\varepsilon).$$

And again using (11.11) and the definition (11.15) of $\mathbf{u}$ we get

$$\tilde{\mathbf{z}}'(t_* + \tau/n) - (\mathbf{w}_n^\sharp)'(0) + \mathbf{w}_n'(0) = \mathbf{u} + O(\varepsilon).$$

Thus we find $g_n = \theta + O(\varepsilon)$ for $\theta = \mathbf{d}f(\mathbf{z}_*, t_*) \cdot \mathbf{u}$.

Now by Lemma 11.2 we know $\theta \ne 0$. Thus if ε is small enough we have $|g_n| \ge \frac{1}{2}|\theta|$, and now (11.19) and (11.20) yield for $|\zeta| = \rho \le 1/2$ the inequality

$$|F_n(\zeta) - F_n(0)| - |F_n(0)| \ge \tfrac{1}{2}|\theta|\rho - O(\varepsilon) - O(\rho^2).$$

Here the right-hand side can be made strictly positive by choosing ρ to be a sufficiently large multiple of ε and then again ε small enough.

Thus indeed there exists ζ_n with $F_n(\zeta_n) = 0$, and by the definition of f this means that $n\tilde{\mathbf{z}}(t_n)$ lies on $\tilde{Z}(t_n)$. Exponentiating, we see that $n\{P(\mathbf{c}_n), Q(\mathbf{c}_n)\}$ lies in $V(\mathbf{c}_n)$ for $\mathbf{c}_n = \mathbf{c}(t_n)$. As $n \ge n_0$ this seems at first sight to complete the proof.

But why is $n\{P(\mathbf{c}_n), Q(\mathbf{c}_n)\} \neq 0$? If this were false then $n\tilde{\mathbf{z}}(t_n)$ would be in $\tilde{\Omega}(t_n)$. Then by (11.18) the point $\mathbf{z}_*$ would be within $O(\varepsilon)$ of a period of $\tilde{\Omega}(t_n)$. Writing this period as an integral linear combination of $\tilde{\mathbf{f}}(t_n), \tilde{\mathbf{g}}(t_n), \tilde{\mathbf{k}}(t_n), \tilde{\mathbf{l}}(t_n)$ we see easily that the coefficients are $O(1)$. It follows that $\mathbf{z}_*$ is within $O(\varepsilon)$ of the corresponding linear combination of $\tilde{\mathbf{f}}(t_*), \tilde{\mathbf{g}}(t_*), \tilde{\mathbf{k}}(t_*), \tilde{\mathbf{l}}(t_*)$. But if ε is small enough, this contradicts the choice of $\mathbf{z}_*$ in Lemma 11.2.

And also why do we get infinitely many different t_n as n varies, as required in (11.1)? Simply because in (11.17) we had $|\tau| \geq 1$ and $\zeta_n = O(\varepsilon)$, so τ/n dominates if ε is sufficiently small.

And finally why is $\mathbf{u} \neq 0$ in (11.15)? Well, the x, y, u, v in (11.10) are real-analytic functions $x(t), y(t), u(t), v(t)$ at $t = t_*$, and (11.15) and (11.13) give $\mathbf{u} = \mathbf{u}(t_*)$ for

$$\mathbf{u}(t) = \tilde{\mathbf{z}}'(t) - x(t)\tilde{\mathbf{f}}'(t) - y(t)\tilde{\mathbf{g}}'(t) - u(t)\tilde{\mathbf{k}}'(t) - v(t)\tilde{\mathbf{l}}'(t).$$

Of course here

$$\tilde{\mathbf{z}}(t) = x(t)\tilde{\mathbf{f}}(t) + y(t)\tilde{\mathbf{g}}(t) + u(t)\tilde{\mathbf{k}}(t) + v(t)\tilde{\mathbf{l}}(t). \tag{11.21}$$

If now by some bad luck $\mathbf{u}(t_*) = 0$ then we just move t_* slightly. We can do this provided $\mathbf{u}(t)$ is not identically zero. But if it were, then from (11.21) we would deduce

$$\delta(x(t))\tilde{\mathbf{f}}(t) + \delta(y(t))\tilde{\mathbf{g}}(t) + \delta(u(t))\tilde{\mathbf{k}}(t) + \delta(v(t))\tilde{\mathbf{l}}(t) = 0,$$

where δ denotes the derivative with respect to either the real or the imaginary part of t. From this it would follow that the real coefficients $\delta(x(t)), \delta(y(t)), \delta(u(t)), \delta(v(t))$ are zero, and so $x(t), y(t), u(t), v(t)$ would be constant. But then (11.21) would contradict Lemma 5.1 on algebraic independence. This really does finish the proof of Theorem P2 in the slightly stronger form with A of arbitrarily large degree.

If we use clusterpoints of $\mathbb{N}\mathbf{z}(\mathbf{c}_*)$ other than 0, then the argument proves more about the set of integers n such that some $\mathbf{c}$ exists. We leave it to the interested reader to explore what can be extracted from the proof.

As anticipated, we now start on the proof that (11.4) is essentially the only example of (11.1), even when A, B, c', c are defined over the algebraic closure $\overline{\mathbb{C}(t)}$, that is, up to multiplication by non-zero elements of this field.

Lemma 11.3. *Suppose that φ, ψ are in $\mathbb{C}(t)$ with $\psi^2 = \varphi^6 + \varphi + t$. Then $\varphi = -t$ and $\psi = \pm t^3$.*

Proof. We easily deduce $\varphi = U/W$, $\psi = V/W^3$ for U, V, W in $\mathbb{C}[t]$ with both U, W and V, W coprime. Thus

$$V^2 = U^6 + (U + tW)W^5. \tag{11.22}$$

Now if $\deg U \leq \deg W$ then tW^6 would dominate on the right-hand side of (11.22), incompatible with the left-hand side V^2. So $\deg U \geq 1 + \deg W$. Now U^6 dominates on the right, and so $\deg V = 3\deg U \geq 3 + 3\deg W$. Thus the maximal degree of the three terms in (11.22) is $N = 2\deg V$.

If U, V are coprime then we can apply abc to (11.22). The number of distinct zeros is at most

$$\deg V + \deg U + \deg(U + tW) + \deg W \le \tfrac{1}{2}N + \tfrac{1}{3}N + \tfrac{1}{3}(N/2 - 3) < N,$$

a contradiction.

Thus we can assume that U, V have a common factor, which must be t up to units. So t does not divide W. Writing $U = tU_1$, $V = tV_1$ we deduce from (11.22) that

$$V_1^2 = t^4 U_1^6 + \frac{U_1 + W}{t} W^5,$$

where $(U_1 + W)/t$ must be a polynomial. Now $\deg V_1 = 2 + 3\deg U_1 \ge 2 + 3\deg W$. Certainly U_1, V_1 are coprime; and if further t, V_1 are coprime, then we can again apply abc. With now $N_1 = 2\deg V_1$ as the maximal degree we find that the number of distinct zeros is at most

$$\deg V_1 + (1 + \deg U_1) + (\deg(U_1 + W) - 1) + \deg W \le \tfrac{1}{2}N_1 + \tfrac{2}{3}(N_1/2 - 2) + \tfrac{1}{3}(N_1/2 - 2) < N_1,$$

another contradiction.

Next suppose t divides V_1 but t^2 does not, so $V_1 = tV_2$ with t, V_2 coprime and

$$V_2^2 = t^2 U_1^6 + \frac{U_1 + W}{t^3} W^5$$

where again $(U_1 + W)/t^3$ must be a polynomial. Now $\deg V_2 = 1 + 3\deg U_1 \ge 1 + 3\deg W$, and with $N_2 = 2\deg V_2$ the zero count is

$$\deg V_2 + (1 + \deg U_1) + (\deg(U_1 + W) - 3) + \deg W \le \tfrac{1}{2}N_2 + \tfrac{2}{3}(N_2/2 - 4) + \tfrac{1}{3}(N_2/2 - 1) < N_2,$$

yet another contradiction.

Finally, if t^2 divides V_1 so $V_1 = t^2 V_3$ then

$$V_3^2 = U_1^6 + \frac{U_1 + W}{t^5} W^5$$

with coprime terms and $\deg V_3 = 3\deg U_1 \ge 3\deg W$, and with $N_3 = 2\deg V_3$ the zero count is

$$\deg V_3 + \deg U_1 + (\deg(U_1 + W) - 5) + \deg W \le \tfrac{1}{2}N_3 + \tfrac{2}{3}(N_3/2 - 15/2) + \tfrac{1}{3}(N_2/2) < N_2,$$

apparently yet another contradiction. But now there is a way out: all the terms could be constant (this was not possible for the first three applications of abc). But then U_1, W would be constant. As $U_1 + W$ is divisible by t, it must vanish. This leads back to $\varphi = -t$, $\psi = \pm t^3$, and the present lemma is proved.

Now we can prove indeed that (11.1) must be essentially (11.4). Clearly $A \neq 0$. By the generic insolvability of the Pell equation over $\overline{\mathbb{C}(t)}$ proved in Section 10 just after (10.2), we can assume that $c' \neq 0$.

In fact by [PT, Proposition 3.6, p. 161] with $g = 2$ the quotient A/B is a convergent in the continued fraction expansion of (10.3) over $\mathbb{C}(t)$. Thus we can suppose that A, B (and so c', c) are over $\mathbb{C}(t)$ (and even $\mathbb{Q}(t)$ but we will not use this). Substituting $X = -c/c'$ we obtain φ, ψ in $\mathbb{C}(t)$ with $\psi^2 = \varphi^6 + \varphi + t$. By Lemma 11.3 we deduce $\varphi = -t$. Thus $c'X + c = c'(X + t)$.

Now we go back to Lemma 11.1 with $K = \overline{\mathbb{C}(t)}$. By (i) the function $A(X) + YB(X)$ has for some n a zero of order $n - 1$ at ∞^+, a pole of order n at ∞^-, and one other zero $\gamma^+ \neq \infty^+$ at which $c'(X + t)$ also vanishes. So $\gamma^+ = (-t, \pm t^3)$; and by changing the sign of B we may suppose that $\gamma^+ = (-t, t^3)$.

Also $X^3 + Y$ has a zero of order 5 at ∞^+, a pole of order 6 at ∞^-, and one other zero at this γ^+. So $g = \frac{A(X)+YB(X)}{X^3+Y}$ has a zero of order $n - 6$ at ∞^+, a pole of order $n-6$ at ∞^-, and no other zeros or poles. By Lemma 10.1(ii) this forces $n = 6$ because of generic Pell insolvability. Thus g must be constant, showing indeed that A, B are constant multiples of X^3, 1 as claimed. The above sign change means that also X^3, -1 turns up.

12. Further remarks

We close this paper with more comments on "likely intersections".

We have shown in Theorem P2 the infinitude of the set T of complex t for which (11.1) is solvable for some A, B in $\mathbb{C}[X]$ with the degree of A not 3 and some $c' \neq 0$, c in $\mathbb{C}$. Equivalently, it is the set of complex t for which there exists a non-negative integer $n \neq 6$ such that

$$np(t) \text{ lies in } \hat{V}(t), \tag{12.1}$$

where $p(t)$ is a certain point on the Jacobian of a certain hyperelliptic curve and $\hat{V}(t)$ is a certain embedding of the curve in the Jacobian (omitting the origin), all depending algebraically on t. This is an analogue of the second sentence of the very first paper [BMZ1] on the subject of likely and unlikely intersections. There it was easy to see that there are infinitely many t for which there exist r, s in $\mathbb{Z}$, not both zero, with

$$t^r(1 - t)^s = 1; \tag{12.2}$$

so easy, in fact, that we did not say how. And without much difficulty we went further to determine some structure and found explicitly the t in $\mathbb{Z}$ and the t in $\mathbb{Q}$; also we considered the t in some fixed number field using Faltings's Theorem, and noted the "sparseness" of the t with fixed degree over $\mathbb{Q}$ using a general result of the first author [M1].

Surprisingly, none of these structure results seems to be clear for our present infinite set T. It is at least obvious that T is in $\overline{\mathbb{Q}}$. But it is not at all obvious even that $T \neq \overline{\mathbb{Q}}$!

In this connection we may note that for any specific $t = t_*$ (apart from those in the finite set for which the Pell equation is solvable) there are at most finitely many A, B, c', c in (11.1) up to proportionality. For otherwise by (12.1) we would get infinitely many

points in $V(t_*)$ defined over $\mathbb{Q}(t_*)$, also contradicting Faltings's Theorem. Now these points have a special cyclic group structure, and for such small rank, simpler results of Chabauty may suffice for the same finiteness conclusion.

These results of Chabauty are proved with p-adic methods, and so we asked Victor Flynn if perhaps similar arguments could be applied with varying t_*, maybe p-adically constrained. He replied very quickly in the affirmative; he shows for example that no non-zero element of $7\mathbb{Z}$ is in T (by (1.6) and (11.2) we see that $t = 0$ is in T). His work appears in the Appendix.

The first main result of [BMZ1] implies that the t in (12.2) have absolute height bounded from above. Thus for us the next natural question is whether this holds for the elements of T. It is of course obvious for the very simplest problem $t^n = 1$; but already for $n(2, \sqrt{2(2-t)}) = 0$ it requires Silverman's Theorem, unfortunately not applicable to our T or (12.1).

One may ask several related questions, for example: does T contain only finitely many roots of unity?

Acknowledgments. We heartily thank both Daniel Bertrand for his interest in and help on these matters and Victor Flynn for allowing us to include his work. We are also grateful to Olaf Merkert for the finite field calculations.

References

[Ab] Abel, N. H.: Ueber die Integration der Differential-Formel $\rho dx/\sqrt{R}$, wenn R und ρ ganze Functionen sind. J. Reine Angew. Math. **1**, 185–221 (1826) ERAM 001.0021cj MR 1577609

[AR] Adams, W. W., Razar, M.: Multiples of points on elliptic curves and continued fractions. Proc. London Math. Soc. **41**, 481–498 (1980) Zbl 0403.14002 MR 0591651

[An] André, Y.: Mumford–Tate groups of mixed Hodge structures and the theorem of the fixed part. Compos. Math. **82**, 1–24 (1992) Zbl 0770.14003 MR 1154159

[BD] Baker, M., DeMarco, L.: Preperiodic points and unlikely intersections. Duke Math. J. **159**, 1–29 (2011) Zbl 1242.37062 MR 2817647

[Berr] Berry, T. G.: On periodicity of continued fractions in hyperelliptic function fields. Arch. Math. (Basel) **55**, 259–266 (1990) Zbl 0728.14027 MR 1075050

[Bert1] Bertrand, D.: Théories de Galois différentielles et transcendance. Ann. Inst. Fourier (Grenoble) **59**, 2773–2803 (2009) Zbl 1226.12002 MR 2649338

[Bert2] Bertrand, D.: Galois descent in Galois theories. In: Arithmetic and Galois Theory of Differential Equations, L. Di Vizio and T. Rivoal (eds.), Sém. Congrès 23, Soc. Math. France, 1–24 (2011) Zbl 06308135 MR 3076077

[Bert3] Bertrand, D.: Special points and Poincaré bi-extensions; with an appendix by Bas Edixhoven. arXiv:1104.5178v1 (2011)

[BMPZ] Bertrand, D., Masser, D., Pillay, A., Zannier, U.: Relative Manin–Mumford for semi-abelian surfaces. Submitted

[BC] Bombieri, E., Cohen, P.: Siegel's lemma, Padé approximations and jacobians. Ann. Scuola Norm. Sup. Pisa **25**, 155–178 (1977) Zbl 1073.11518 MR 1655513

[BMZ1] Bombieri, E., Masser, D., Zannier, U.: Intersecting a curve with algebraic subgroups of multiplicative groups. Int. Math. Res. Notices **1999**, no. 20, 1119–1140 Zbl 0938.11031 MR 1728021

[BMZ2] Bombieri, E., Masser, D., Zannier, U.: Finiteness results for multiplicatively dependent points on complex curves. Michigan Math. J. **51**, 451–466 (2003) Zbl 1048.11056 MR 2021000

[BMZ3] Bombieri, E., Masser, D., Zannier, U.: On unlikely intersections of complex varieties with tori. Acta Arith. **133**, 309–323 (2008) Zbl 1162.11031 MR 2457263

[CF] Cassels, J. W. S., Flynn, E. V.: Prolegomena to a Middlebrow Arithmetic of Curves of Genus 2, London Math. Soc. Lecture Note Ser. 230, Cambridge Univ. Press (1996) Zbl 0857.14018 MR 1406090

[C1] Tchebichef, P. [= Chebyshev, P.]: Sur l'intégration des différentielles qui contiennent une racine carrée d'un polynôme du troisième ou du quatrième degré. J. Math. Pures Appl. **2**, 1–42 (1857)

[C2] Tchébichef, P. [= Chebyshev, P.]: Sur l'intégration de la différentielle $\frac{x+A}{\sqrt{x^4+\alpha x^3+\beta x^2+\gamma x+\delta}}dx$. J. Math. Pures Appl. **9**, 225–241 (1864)

[CMZ] Corvaja, P., Masser, D., Zannier, U.: Sharpening 'Manin–Mumford' for certain algebraic groups of dimension 2. Enseign. Math. **59**, 225–269 (2013) MR 3189035

[Dave] Davenport, J. H.: On the Integration of Algebraic Functions. Lecture Notes in Computer Sci. 102, Springer (1981) Zbl 0471.14009 MR 0617377

[Davi] David, S.: Fonctions theta et points de torsion des variétés abéliennes. Compos. Math. **78**, 121–160 (1991) Zbl 0741.14025 MR 1104784

[EEHK] Ellenberg, J., Elsholtz, C., Hall, C., Kowalski, E.: Non-simple abelian varieties in a family: geometric and analytic approaches. J. London Math. Soc. **80**, 135–154 (2009) Zbl 1263.11064 MR 2520382

[EHK] Ellenberg, J., Hall, C., Kowalski, E.: Expander graphs, gonality and variation of Galois representations. Duke Math. J. **161**, 1233–1275 (2012) Zbl 1262.14021 MR 2922374

[G] Grant, D.: Formal groups in genus two. J. Reine Angew. Math. **411**, 96–121 (1990) Zbl 0702.14025 MR 1072975

[H] Halphen, G.-H.: Fonctions elliptiques II. Gauthier-Villars, Paris (1888)

[HL] Hellegouarch, Y., Lozach, M.: Équation de Pell et points d'ordre fini. In: Analytic and Elementary Number Theory (Marseille, 1983), Publ. Math. Orsay 86, 72–95 (1986) Zbl 0594.10007 MR 0844586

[K1] Katz, N. M.: Monodromy of families of curves: applications of some results of Davenport-Lewis. In: Séminaire de Théorie de Nombres, Paris 1979–80, M.-J. Bertin (ed.), Progr. Math. 12, Birkhäuser, Boston, 171–195 (1981) Zbl 0475.14025 MR 0633896

[K2] Katz, N. M.: Affine cohomological transforms, perversity, and monodromy. J. Amer. Math. Soc. **6**, 149–222 (1993) Zbl 0815.14011 MR 1161307

[L] Lang, S.: Complex Analysis. Addison-Wesley (1977) Zbl 0366.30001 MR 0477000

[LB] Lange, H., Birkenhake, C.: Complex Abelian Varieties. Springer (1992) Zbl 0779.14012 MR 1217487

[M1] Masser, D.: Specializations of finitely generated subgroups of abelian varieties. Trans. Amer. Math. Soc. **311**, 413–424 (1989) Zbl 0673.14016 MR 0974783

[M2] Masser, D.: Specializations of endomorphism rings of abelian varieties. Bull. Soc. Math. France **124**, 457–476 (1996) Zbl 0866.11040 MR 1415735

[M3] Masser, D.: Specializations of some hyperelliptic Jacobians. In: Number Theory in Progress, K. Győry et al. (eds.), de Gruyter, Berlin, 293–307 (1999) Zbl 0942.14015 MR 1689511

[MW1] Masser, D., Wüstholz, G.: Periods and minimal abelian subvarieties. Ann. of Math. **137**, 407–458 (1993) Zbl 0796.11023 MR 1207211

[MW2] Masser, D., Wüstholz, G.: Factorization estimates for abelian varieties. Publ. Math. IHES **81**, 5–24 (1995) Zbl 0854.11030 MR 1361754

[MZ1] Masser, D., Zannier, U.: Torsion anomalous points and families of elliptic curves. C. R. Math. Acad. Sci. Paris **346**, 491–494 (2008) Zbl 1197.11066 MR 2412783

[MZ2] Masser, D., Zannier, U.: Torsion anomalous points and families of elliptic curves. Amer. J. Math. **132**, 1677–1691 (2010) Zbl 1225.11078 MR 2766181

[MZ3] Masser, D., Zannier, U.: Torsion points on families of squares of elliptic curves. Math. Ann. **352**, 453–484 (2012) Zbl 1306.11047 MR 2874963

[MZ4] Masser, D., Zannier, U.: Torsion points on families of products of elliptic curves. Adv. Math. **259**, 116–133 (2014) Zbl 06296246 MR 3197654

[Mc] McMullen, C. T.: Teichmüller curves in genus two: Torsion divisors and ratios of sines. Invent. Math. **165**, 651–672 (2006) Zbl 1103.14014 MR 2242630

[N] Nadel, A.: The nonexistence of certain level structures on abelian varieties over complex function fields. Ann. of Math. **129**, 161–178 (1989) Zbl 0675.14018 MR 0979604

[P] Paysant-Le Roux, R.: Périodicité des fractions continues dans un corps de fonctions hyperelliptiques. Arch. Math. (Basel) **61**, 46–58 (1993) Zbl 0778.11065 MR 1222088

[Pil] Pila, J.: Integer points on the dilation of a subanalytic surface. Quart. J. Math. **55**, 207–223 (2004) Zbl 1111.32004 MR 2068319

[PZ] Pila, J., Zannier, U.: Rational points in periodic analytic sets and the Manin–Mumford conjecture. Rend. Lincei Mat. Appl. **19**, 149–162 (2008) Zbl 1164.11029 MR 2411018

[Pin] Pink, R.: A common generalization of the conjectures of André–Oort, Manin–Mumford, and Mordell–Lang. Manuscript dated 17th April 2005 (13 pages)

[PT] van der Poorten, A. J., Tran, X. C.: Quasi-elliptic integrals and periodic continued fractions. Monatsh. Math. **131**, 155–169 (2000) Zbl 0972.11062 MR 1798560

[Schi] Schinzel, A.: On some problems of the arithmetical theory of continued fractions II. Acta Arith. **7**, 287–298 (1962) Zbl 0112.28001 MR 0139566

[Schm] Schmidt, W. M.: On continued fractions and diophantine approximation in power series fields. Acta Arith. **95**, 139–166 (2000) Zbl 0987.11041 MR 1785412

[ST] Serre, J-P., Tate, J.: Good reduction of abelian varieties. Ann. of Math. **88**, 492–517 (1968) Zbl 0172.46101 MR 0236190

[Si1] Silverman, J. H.: Heights and the specialization map for families of abelian varieties. J. Reine Angew. Math. **342**, 197–211 (1983) Zbl 0505.14035 MR 0703488

[Si2] Silverman, J. H.: The Arithmetic of Elliptic Curves. Springer (1986) Zbl 0596.00007 MR 0817210

[St] Stoll, M.: Two simple 2-dimensional abelian varieties defined over $\mathbb{Q}$ with Mordell–Weil rank at least 19. C. R. Acad. Sci. Paris Sér. I **321**, 1341–1344 (1995) Zbl 0859.11033 MR 1363577

[Za] Zannier, U.: Some Problems of Unlikely Intersections in Arithmetic and Geometry. Ann. of Math. Stud. 181, Princeton Univ. Press (2012) Zbl 1246.14003 MR 2918151

[Zh] Zhang, S.: Small points and Arakelov theory. Documenta Math. (Extra Volume II of ICM 1998), 217–225. Zbl 0912.14008 MR 1648072

[Zi] Zilber, B.: Exponential sums equations and the Schanuel conjecture. J. London Math. Soc. **65**, 27–44 (2002) Zbl 1030.11073 MR 1875133

Appendix (by E. V. Flynn): An application of Chabauty's theorem to a family of curves

In this appendix, we shall consider the family of genus 2 curves

$$\mathcal{H}_t : Y^2 = X^6 + X + t, \tag{1}$$

where $t \in \mathbb{C}$ is such that $X^6 + X + t$ has only simple roots. Since the discriminant of this sextic with respect to X is $3125 - 46656t^5$, this condition is equivalent to requiring that t avoids any 5th root of $3125/46656$, and in particular the condition is satisfied by any $t \in \mathbb{Q}$. Any such $\mathcal{H}_t$ is of genus 2; let J_t denote the Jacobian of $\mathcal{H}_t$. The above curve is defined over $K = \mathbb{Q}(t)$. Let ∞^+, ∞^- denote the points on the non-singular curve that lie over the singular point at infinity, which should be regarded as members of $\mathcal{H}_t(K)$, since the coefficient of X^6 is in $(K^*)^2$. We shall adopt the customary shorthand notation $\{P_1, P_2\}$ to denote the divisor class $[P_1 + P_2 - \infty^+ - \infty^-]$, which is in $J_t(K)$ when P_1, P_2 are points on $\mathcal{H}_t$ and either P_1, P_2 are both K-rational, or they are quadratic over K and conjugate. Consider the following embeddings:

$$\begin{aligned} \mu &: \mathcal{H}_t(K) \hookrightarrow J_t(K) : P \mapsto [P - \infty^+] = \{P, \infty^-\}, \\ \mu' &: \mathcal{H}_t(K) \hookrightarrow J_t(K) : P \mapsto [P - \infty^-] = \{P, \infty^+\}. \end{aligned} \tag{2}$$

Further define

$$q_t := [\infty^- - \infty^+] = \{\infty^-, \infty^-\} \in J_t(K), \tag{3}$$

which is in the image of μ. Making use of the divisor of the function $Y + X^3$, it is straightforward to compute $2q_t$, and then all nq_t for $-3 \le n \le 3$:

$$\begin{aligned} -3q_t &= \{(-t, t^3), \infty^+\}, \quad -2q_t = \{(-t, t^3), \infty^-\}, \quad -1q_t = \{\infty^+, \infty^+\}, \\ 0q_t &= \{\infty^+, \infty^-\}, \\ 1q_t &= \{\infty^-, \infty^-\}, \quad 2q_t = \{(-t, -t^3), \infty^+\}, \quad 3q_t = \{(-t, -t^3), \infty^-\}. \end{aligned} \tag{4}$$

Clearly $nq_t \in \operatorname{im}\mu$ for $n = -2, 0, 1, 3$ and $nq_t \in \operatorname{im}\mu'$ for $n = -3, -1, 0, 2$.

We are interested in finding sufficient conditions on t (which we wish to include infinitely many $t \in \mathbb{Q}$) such that $nq_t \notin \operatorname{im}\mu$ for $|n| > 3$. It turns out to be more elegant to rephrase this as: $nq_t \notin \operatorname{im}\mu \cup \operatorname{im}\mu'$ for $|n| > 3$.

When t is algebraic over $\mathbb{Q}$, so that K is a number field, this is a problem that should be amenable to constructive Chabauty techniques, which provide explicit bounds on the order of the intersection of an embedding of a curve $\mathcal{C}$ of genus g into its Jacobian J, and a rank r subgroup of the Mordell–Weil group $J(K)$, provided that $r < g$. In this case, the genus of $\mathcal{H}_t$ is $g = 2$ and we are trying to find the intersection of $\mu(\mathcal{H}_t(K))$ with the rank 1 subgroup of $J_t(K)$ generated by q_t. There is a substantial literature on applications of Chabauty techniques, which we shall not attempt to list here; for genus 2, there are different styles used to find explicit bounds, such as those in [1], [2], [4], [5], [6], and there is an implementation for numerical genus 2 examples, due to Michael Stoll, in Magma [8] (see also [3] for the original article of Chabauty). We shall follow the methodology of [6]

and [2, Chapter 13] using explicit local parameters to find power series over $\mathbb{Z}_p$. We shall give the details below, and have provided a Maple file mzf.map at [7] which checks all of the following steps.

The first step towards applying these techniques is to find a multiple of q_t which is in the kernel of the reduction map modulo some prime p. Using the multiples of q_t in (2) and adding $2q_t + 3q_t$, we find

$$E_t := 5q_t = \{(-t, -t^3), (-t, -t^3)\}. \tag{5}$$

We now impose the condition that

$$\begin{aligned} &t = t_0 \neq 0, \text{ where } t_0 \text{ is algebraic over } \mathbb{Q} \text{ and there exists } p > 5 \\ &\text{and an embedding of } \mathbb{Q}(t_0) \text{ into } \mathbb{Q}_p \text{ with } |t_0|_p < 1. \end{aligned} \tag{6}$$

We represent this by setting

$$t = t_0 = u_0 p^k, \quad \text{where} \quad |u_0|_p = 1 \text{ and } k \geq 1. \tag{7}$$

This forces E_t to be in the kernel of reduction modulo p. Note that this condition on t_0 includes all $a/b \in \mathbb{Q}$ with $a \neq 0$, $b \in \mathbb{Z}$ and $\mathrm{hcf}(a, b) = 1$ such that there exists a prime $p > 5$ with $p \mid a$ (which in turn include all members of $\mathbb{Z}$ outside a set of density 0). It also includes algebraic numbers t_0 of arbitrary degree over $\mathbb{Q}$ such as, for any $k \in \mathbb{N}$ not divisible by 11, $t_0 = 12^{1/k} - 1$, which is of degree k over $\mathbb{Q}$ and for which there is an embedding into $\mathbb{Q}_{11}$ with $|t_0|_{11} < 1$.

What is somewhat surprising is that we shall find condition (6), which merely places E_t in the kernel of the reduction map modulo p, already to be sufficient to give our desired result about multiples of q_{t_0} using a p-adic Chabauty technique, when normally one might expect further congruence conditions to be required. Although there are a number of worked examples using these techniques in the literature, we shall nevertheless give an outline of the details here, on the grounds that the naive bound is insufficient, and so there is a finesse required towards the end of the argument, for which it is helpful to see the actual power series.

Recall from [2, Chapter 2] that for a general curve of genus 2,

$$Y^2 = f_6X^6 + f_5X^5 + f_4X^4 + f_3X^3 + f_2X^2 + f_1X + f_0, \tag{8}$$

there is an embedding of the Jacobian variety into $\mathbb{P}^{15}$, where $D = \{(x_1, y_1), (x_2, y_2)\}$ is mapped to $\mathbf{a} = (a_0, \ldots, a_{15})$ with

$$\begin{aligned} &a_{15} = (x_1 - x_2)^2, \quad a_{14} = 1, \quad a_{13} = x_1 + x_2, \quad a_{12} = x_1x_2, \quad a_{11} = x_1x_2(x_1 + x_2), \\ &a_{10} = (x_1x_2)^2, \quad a_9 = (y_1 - y_2)/(x_1 - x_2), \quad a_8 = (x_2y_1 - x_1y_2)/(x_1 - x_2), \\ &a_7 = (x_2^2y_1 - x_1^2y_2)/(x_1 - x_2), \quad a_6 = (x_2^3y_1 - x_1^3y_2)/(x_1 - x_2), \\ &a_5 = (F_0(x_1, x_2) - 2y_1y_2)/(x_1 - x_2)^2, \\ &a_4 = (F_1(x_1, x_2) - (x_1 + x_2)y_1y_2)/(x_1 - x_2)^2, \\ &a_3 = (x_1x_2)a_5, \quad a_2 = (G_0(x_1, x_2)y_1 - G_0(x_2, x_1)y_2)/(x_1 - x_2)^3, \\ &a_1 = (G_1(x_1, x_2)y_1 - G_1(x_2, x_1)y_2)/(x_1 - x_2)^3, \quad a_0 = a_5^2, \end{aligned} \tag{9}$$

where

$$\begin{aligned}
F_0(x_1,x_2) &= 2f_0 + f_1(x_1+x_2) + 2f_2(x_1x_2) + f_3(x_1x_2)(x_1+x_2)\\
&\quad + 2f_4(x_1x_2)^2 + f_5(x_1x_2)^2(x_1+x_2) + 2f_6(x_1x_2)^3,\\
F_1(x_1,x_2) &= f_0(x_1+x_2) + 2f_1(x_1x_2) + f_2(x_1x_2)(x_1+x_2) + 2f_3(x_1x_2)^2\\
&\quad + f_4(x_1x_2)^2(x_1+x_2) + 2f_5(x_1x_2)^3 + f_6(x_1x_2)^3(x_1+x_2),\\
G_0(x_1,x_2) &= 4f_0 + f_1(x_1+3x_2) + f_2(2x_1x_2+2x_2^2) + f_3(3x_1x_2^2+x_2^3)\\
&\quad + 4f_4(x_1x_2^3) + f_5(x_1^2x_2^3+3x_1x_2^4) + f_6(2x_1^2x_2^4+2x_1x_2^5),\\
G_1(x_1,x_2) &= f_0(2x_1+2x_2) + f_1(3x_1x_2+x_2^2) + 4f_2(x_1x_2^2) + f_3(x_1^2x_2^2+3x_1x_2^3)\\
&\quad + f_4(2x_1^2x_2^3+2x_1x_2^4) + f_5(3x_1^2x_2^4+x_1x_2^5) + 4f_6(x_1^2x_2^5).
\end{aligned} \tag{10}$$

With respect to this embedding,

$$\begin{aligned}
0q_t &= [1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],\\
1q_t &= [0,0,0,0,0,0,0,0,-1,0,0,1,0,0,0,0,0],\\
2q_t &= [0,0,0,0,0,0,0,0,t^2,-t,1,t^2,-t,0,0,0,1],\\
E_t = 5q_t &= \Bigg[\frac{(12t^5-1)^2}{16t^{12}}, \frac{40t^{10}-16t^5+1}{8t^8}, \frac{-(56t^{10}-18t^5+1)}{8t^9},\\
&\quad \frac{-(12t^5-1)}{4t^4}, \frac{10t^5-1}{4t^5}, \frac{-(12t^5-1)}{4t^6}, \frac{1}{2}, \frac{2t^5-1}{2t},\\
&\quad \frac{-(4t^5-1)}{2t^2}, \frac{6t^5-1}{2t^3}, t^4, -2t^3, t^2, -2t, 1, 0\Bigg].
\end{aligned} \tag{11}$$

We recall from [2, Chapter 7] that $s_1 = a_1/a_0$ and $s_2 = a_2/a_0$ give a pair of local parameters; there is a formal group law, defined over $\mathbb{Z}[f_0,\ldots,f_6]$, and the formal logarithm power series $\log_1$, $\log_2$ and formal exponential power series $\exp_1$, $\exp_2$ (available at local/log and local/exp in [7]), up to terms of total degree 5 in s_1, s_2 are, when specialised to our curve $\mathcal{H}_t$ in (1):

$$\begin{aligned}
\log_1 &= s_1 + \tfrac{1}{3}s_2^3 + 12ts_1^3s_2^2 + 5s_1^4s_2 + \text{terms of degree } \geq 7,\\
\log_2 &= s_2 + 4s_1^3s_2^2 + 12ts_1^2s_2^3 + \text{terms of degree } \geq 7,\\
\exp_1 &= s_1 - \tfrac{1}{3}s_2^3 - 5s_1^4s_2 - 12ts_1^3s_2^2 + \text{terms of degree } \geq 7,\\
\exp_2 &= s_2 - 4s_1^3s_2^2 - 12ts_1^2s_2^3 + \text{terms of degree } \geq 7.
\end{aligned} \tag{12}$$

For each of these power series, the denominator occurring in any term of total degree n divides $n!$. Computing the local parameters $s_1(E_t) = a_1/a_0$ and $s_2(E_t) = a_2/a_0$ for E_t in (11) and using the condition on t in (7), we see that

$$\begin{aligned} s_1(E_t) &= 2t^4(40t^{10} - 16t^5 + 1)/(12t^5 - 1)^2 \\ &= 2u_0^4 p^{4k}(1 + 8u_0^5 p^{5k} + 88u_0^{10} p^{10k} + O(p^{15k})), \\ s_2(E_t) &= -2t^3(56t^{10} - 18t^5 + 1)/(12t^5 - 1)^2 \\ &= -2u_0^3 p^{3k}(1 + 6u_0^5 p^{5k} + 56u_0^{10} p^{10k} + O(p^{15k})), \end{aligned} \tag{13}$$

where $O(p^r)$ denotes up^r for some $u \in \mathbb{Z}_p$ with $|u|_p \le 1$. Applying $\exp(m \log E_t)$, by combining (12), (13) and taking account of the denominators in (12) gives the following pair of local parameters for mE:

$$\begin{aligned} s_1(mE_t) &= \tfrac{2}{3} m u_0^4 p^{4k}(3 + 4m^2 u_0^5 p^{5k} + 20u_0^5 p^{5k} \\ &\quad + 72m^2 u_0^{10} p^{10k} + 192u_0^{10} p^{10k} + O(p^{15k})), \\ s_2(mE_t) &= -2m u_0^3 p^{3k}(1 + 6u_0^5 p^{5k} + 56u_0^{10} p^{10k} + O(p^{15k})). \end{aligned} \tag{14}$$

Using the local power series local/local.coordinates at [7] (and described in [2, Chapter 7]), we can find the $\mathbb{P}^{15}$ embedding of any point in the kernel of reduction, given a pair of local parameters; substituting (14) into these gives the following $\mathbb{P}^{15}$ embedding $[a_0(mE_t), \ldots, a_{15}(mE_t)]$ for mE_t:

$$\begin{aligned} a_0(mE_t) &= 1, \\ a_1(mE_t) &= \tfrac{2}{3} m u_0^4 p^{4k}(3 + 4m^2 u_0^5 p^{5k} + 20u_0^5 p^{5k} \\ &\quad + 72m^2 u_0^{10} p^{10k} + 192u_0^{10} p^{10k} + O(p^{15k})), \\ a_2(mE_t) &= -2m u_0^3 p^{3k}(1 + 6u_0^5 p^{5k} + 56u_0^{10} p^{10k} + O(p^{15k})), \\ a_3(mE_t) &= \tfrac{4}{9} m^2 u_0^8 p^{8k}(9 - 12m^2 u_0^5 p^{5k} + 120u_0^5 p^{5k} \\ &\quad - 272m^2 u_0^{10} p^{10k} + 1552u_0^{10} p^{10k} + 16m^4 u_0^{10} p^{10k} + O(p^{15k})), \\ a_4(mE_t) &= -\tfrac{4}{3} m^2 u_0^7 p^{7k}(3 + 4m^2 u_0^5 p^{5k} + 38u_0^5 p^{5k} \\ &\quad + 480u_0^{10} p^{10k} + 96m^2 u_0^{10} p^{10k} + O(p^{15k})), \\ a_5(mE_t) &= -4m^2 u_0^6 p^{6k}(-1 - 12u_0^5 p^{5k} - 148u_0^{10} p^{10k} + 4m^2 u_0^{10} p^{10k} + O(p^{15k})), \\ a_6(mE_t) &= 8m^3 u_0^{12} p^{12k}(1 + 20u_0^5 p^{5k} + 4m^2 u_0^5 p^{5k} + O(p^{15k})), \\ a_7(mE_t) &= -\tfrac{8}{3} m^3 u_0^{11} q^{11k}(3 + 8m^2 u_0^5 p^{5k} + 58u_0^5 p^{5k} + O(p^{10k})), \\ a_8(mE_t) &= \tfrac{8}{3} m^3 u_0^{10} p^{10k}(3 + 56u_0^5 p^{5k} + 4m^2 u_0^5 p^{5k} + O(p^{10k})), \\ a_9(mE_t) &= -8m^3 u_0^9 p^9(1 + 18u_0^5 p^{5k} + O(p^{10k})), \\ a_{10}(mE_t) &= 16m^4 u_0^{16} p^{16k}(1 + O(p^{5k})), \quad a_{11}(mE_t) = 32u_0^{15} p^{15k} m^4(-1 + O(p^{5k})), \\ a_{12}(mE_t) &= 16m^4 u_0^{14} p^{14}(1 + O(p^{5k})), \quad a_{13}(mE_t) = -32u_0^{13} p^{13k} m^4(1 + O(p^{5k})), \\ a_{14}(mE_t) &= 16m^4 u_0^{12} p^{12k}(1 + O(p^{5k})), \quad a_{15}(mE_t) = m^6 u_0^{20} O(p^{20k}). \end{aligned} \tag{15}$$

We also recall, from [2, Chapter 3], that for $D = \{(x_1, y_1), (x_2, y_2)\}$ there is an embedding of the Kummer surface given by (k_1, k_2, k_3, k_4), where

$$k_1 = 1, \quad k_2 = x_1 + x_2, \quad k_3 = x_1 x_2, \quad k_4 = a_5, \tag{16}$$

where a_5 is the function given in (9). We observe that any D in the image of either μ or μ' must have $k_1 = 0$. We also recall from [2, Chapter 3] (available from jacobian.variety/bilinear.forms at [7]) that if $\mathbf{a}, \mathbf{b}$ are on the Jacobian variety, given as members of $\mathbb{P}^{15}$ using the embedding in (9), there are bilinear forms $\phi_{ij}(\mathbf{a}, \mathbf{b})$ which give $k_i(\mathbf{a}-\mathbf{b})k_j(\mathbf{a}+\mathbf{b})$. The bilinear form $\phi_{31}(\mathbf{a}, \mathbf{b})$, when specialised to our curve $\mathcal{H}_t$ in (1), is

$$\begin{aligned}\phi_{31} = &-2a_8b_1 + 2a_7b_2 - 4a_4b_4 + 2b_7a_2 - 2b_8a_1 + b_0a_{12} + a_5b_3 + a_3b_5\\ &+ a_0b_{12} + 2ta_5b_{14} + a_{13}b_5 + a_5b_{13} - 4ta_9b_9 + 2ta_{14}b_5 - 2a_4b_{14}\\ &-2a_8b_9 - 2a_{14}b_4 - 2a_9b_8 + 16tb_{15}a_{12} + 68ta_{12}b_{12} + 4ta_{15}b_{15}\\ &+ 2b_{10}a_3 + 4a_6b_6 + 2a_{10}b_3 + 8b_{11}a_{12} + 8b_{12}a_{11} - 2a_{13}b_{10}\\ &+ 2a_{15}b_{11} - 2a_{10}b_{13} + 2a_{11}b_{15} - 2a_{14}b_{14} - 4ta_{13}b_{11}\\ &+ 16ta_{15}b_{12} - 4ta_{11}b_{13}.\end{aligned} \tag{17}$$

Now define

$$\begin{aligned}\psi_0(m) &= \phi_{31}(0q_t, mE_t) = m^4u_0^{14}p^{14k}(16 + O(p^{5k})),\\ \psi_1(m) &= \phi_{31}(1q_t, mE_t) = mu_0^3p^{3k}(4 + O(p^{5k})),\\ \psi_2(m) &= \phi_{31}(2q_t, mE_t) = \tfrac{8}{3}mu_0^5p^{10k}\big(u_0^5(m+1)(2m+1) + O(p^{5k})\big),\end{aligned} \tag{18}$$

all of which are members of $\mathbb{Z}_p[[m]]$, for which the coefficient of m^r tends to 0 in $\mathbb{Z}_p$ as $r \to \infty$. From the above discussion, we see that, for any $\ell = 0, 1, 2$,

$$\ell q_t + mE_t \in \operatorname{im}\mu \cup \operatorname{im}\mu' \;\Rightarrow\; \psi_\ell(m) = 0. \tag{19}$$

We are now in a position to prove our main result.

Theorem. *Let $\mathcal{H}_t$ be as in* (1)*, let μ, μ' be as given in* (2)*, let q_t be as in* (3) *and let $t = t_0$ satisfy the condition given in* (6)*,* (7) *for some prime $p > 5$. Then for $n \in \mathbb{Z}$ and $n > 3$, we have $nq_t \notin \operatorname{im}\mu \cup \operatorname{im}\mu'$.*

Proof. First note that any nq_t, for $n \in \mathbb{Z}$, must be one of: mE_t, $q_t + mE_t$, $2q_t + mE_t$, $-q_t + mE_t$, $-2q_t + mE_t$, for some $m \in \mathbb{Z}$. We wish to show that $nq_t \in \operatorname{im}\mu \cup \operatorname{im}\mu'$ (for $n \in \mathbb{Z}$) only when $|n| \le 3$. Since $E_t = 5q_t$, and since this condition is invariant under $\pm$, this is equivalent to showing that (for $m \in \mathbb{Z}$): $mE_t \in \operatorname{im}\mu \cup \operatorname{im}\mu'$ only when $m = 0$, $q_t + mE_t \in \operatorname{im}\mu \cup \operatorname{im}\mu'$ only when $m = 0$, and $2q_t + mE_t \in \operatorname{im}\mu \cup \operatorname{im}\mu'$ only when $m = 0, -1$. The first two of these follow immediately from (18), (19), since the power series $16 + O(p^{5k})$, $4 + O(p^{5k})$ each have constant term with $|\ |_p = 1$, which is strictly greater than for the coefficients of all subsequent powers of m, and so there are no further roots $m \in \mathbb{Z}_p$ (and so no further roots $m \in \mathbb{Z}$) of $\psi_0(m)$, $\psi_1(m)$ apart from $m = 0$.

The interesting case is that of $\psi_2(m)$. We know $m = 0, -1$ to be solutions (since $2q_t$, $2q_t - E_t$ are indeed in $\operatorname{im}\mu \cup \operatorname{im}\mu'$), so that

$$\psi_2(m) = \tfrac{8}{3}m(m+1)u_0^5p^{10k}(u_0^5(2m+1) + O(p^{5k})).$$

In this case, the number of possible solutions $m \in \mathbb{Z}_p$ is bounded above by 3, which is strictly greater than the number of known solutions $m = 0, -1 \in \mathbb{Z}$, and the Chabauty bound fails in this case. However, there is a finesse in this case which allows us to identify this third solution $m \in \mathbb{Z}_p$. By Hensel's Lemma (keeping in mind that $t = t_0$ satisfies the condition (6), (7)), there is a root $w_t \in \mathbb{Z}_p$ of the sextic $X^6 + X + t$ with $w_t \equiv -t \pmod{p}$ and so $D_t := \{(-t, -t^3), (w_t, 0)\}$ is in the kernel of reduction modulo p, and satisfies $2D_t = E_t$. This can be regarded as $\frac{1}{2}E_t$ (within the kernel of reduction), in the sense that $m = 1/2$ gives D_t when inserted into $\exp(m \log E_t)$. This means that $2q_t - \frac{1}{2}E_t = \{\infty^+, (-t, -t^3)\} + \{(w, 0), (-t, t^3)\} = \{\infty^+, (w, 0)\} \in \operatorname{im} \mu'$, and so $-1/2 \in \mathbb{Z}_p$ must be a root of $\psi_2(m)$. Therefore $\psi_2(m) = \frac{8}{3}m(m+1)(2m+1)u_0^5 p^{10k}(u_0^5 + O(p^{5k}))$, and the complete list of solutions in $\mathbb{Z}_p$ is given by $m = 0, -1, -1/2$. Therefore the only solutions in $\mathbb{Z}$ are $m = 0, -1$, as required.

Acknowledgments. The author thanks EPSRC for support: grant number EP/F060661/1.

References

[1] Bruin, N.: Chabauty methods and covering techniques applied to generalised Fermat equations. Ph.D. thesis, Univ. Leiden (1999) MR 1916903

[2] Cassels, J. W. S., Flynn, E. V.: Prolegomena to a Middlebrow Arithmetic of Curves of Genus 2. London Math. Soc. Lecture Note Ser. 230, Cambridge Univ. Press, Cambridge (1996) Zbl 0857.14018 MR 1406090

[3] Chabauty, C.: Sur les points rationnels des variétés algébriques dont l'irrégularité est supérieure à la dimension. C. R. Acad. Sci. Paris **212**, 1022–1024 (1941) Zbl 0025.24903 MR 0011005

[4] Coleman, R. F.: Effective Chabauty. Duke Math. J. **52**, 765–780 (1985) Zbl 0588.14015 MR 0808103

[5] Flynn, E. V.: A flexible method for applying Chabauty's theorem. Compos. Math. **105**, 79–94 (1997) Zbl 0882.14009 MR 1436746

[6] Flynn, E. V., Poonen, B., Schaefer, E.: Cycles of quadratic polynomials and rational points on a genus 2 curve. Duke Math. J. **90**, 435–463 (1997) Zbl 0958.11024 MR 1480542

[7] Flynn, E. V.: http://people.maths.ox.ac.uk/flynn/genus2

[8] The Magma Computational Algebra System, http://magma.maths.usyd.edu.au/magma/