Special points on the product of two modular curves.

Bas Edixhoven*

September 1996

1 Introduction.

It is well known that the j-invariant establishes a bijection between $\mathbb{C}$ and the set of isomorphism classes of elliptic curves over $\mathbb{C}$, see for example [10]. The endomorphism ring of an elliptic curve E over $\mathbb{C}$ is either $\mathbb{Z}$ or an order in an imaginary quadratic extension of $\mathbb{Q}$; in the second case E is said to be a CM elliptic curve (CM meaning complex multiplication). A complex number x is said to be CM if the corresponding elliptic curve over $\mathbb{C}$ is CM. A point (x_1, x_2) in $\mathbb{C}^2$ is defined to be CM if both x_1 and x_2 are CM. The aim of this article is to determine all irreducible algebraic curves C in $\mathbb{C}^2$ containing infinitely many CM points. In other words, we want to determine all irreducible polynomials f in $\mathbb{C}[x_1, x_2]$ that vanish at infinitely many CM points. The motivation for doing this comes from a conjecture of Frans Oort (see [7, Chapter IV, §1] for a precise statement), saying roughly that the irreducible components of the Zariski closure of any set of CM points in any Shimura variety are sub Shimura varieties. For the irreducible components of dimension zero this is trivially true. For those of dimension one Oort's conjecture was in fact stated earlier by Yves André as a problem in [2, Chapter X, §1].

We view $\mathbb{C}^2$ as the Shimura variety which is the moduli space of pairs of elliptic curves. Then the irreducible sub Shimura varieties of dimension one are the following: $\mathbb{C} \times \{x_2\}$ with x_2 a CM point, $\{x_1\} \times \mathbb{C}$ with x_1 a CM point, or the image in $\mathbb{C}^2$, under the usual map, of the modular curve $Y_0(n)$ for some integer $n \geq 1$. Recall that, for $n \geq 1$, $Y_0(n)$ is the modular curve classifying elliptic curves with a cyclic subgroup of order n, or, equivalently, cyclic isogenies of degree n between elliptic curves. The usual map from $Y_0(n)$ to $\mathbb{C}^2$ sends an isogeny to its source and target, i.e., $\phi: E_1 \to E_2$ is sent to $(j(E_1), j(E_2))$. We will prove the following result, giving evidence for the conjecture just mentioned.

1.1 Theorem. *Assume the generalized Riemann hypothesis for imaginary quadratic fields. Let C be an irreducible algebraic curve in $\mathbb{C}^2$ containing infinitely many CM points and such that neither of its projections to $\mathbb{C}$ is constant. Then C is the image of $Y_0(n)$ for some $n \geq 1$.*

1.2 Remark. In the proof of Theorem 1.1 we will see that the state of the art in analytic number theory is such that the Riemann hypothesis is "almost not needed" (see Remark 5.4). It is clear

*partially supported by the Institut Universitaire de France

that Theorem 1.1 implies similar statements for curves contained in the product of two modular curves. In particular, if one assumes GRH, Oort's conjecture is true for curves contained in the product of two modular curves. □

1.3 Remark. Ben Moonen has proved Oort's conjecture for the sets of CM points in moduli spaces of abelian varieties such that there exists a prime number p at which all the CM points are canonical in the sense that they have an ordinary reduction of which they are the Serre-Tate canonical lift (see [7, Chapter IV, §1]). Yves André has proved the conclusion of Theorem 1.1 with the Riemann hypothesis replaced by the assumption that the Zariski closure of C in $\mathbb{P}^1 \times \mathbb{P}^1$ meets $\{\infty\} \times \mathbb{C}$ only in points (∞, x_2) with x_2 a CM point (see [1]). In the case where C meets the union of $\{\infty\} \times \mathbb{C}$ and $\mathbb{C} \times \{\infty\}$ only in $\infty \times \infty$ he has a very simple proof. □

The idea of the proof of Theorem 1.1 is the following. We use the Galois action on the set of CM j-invariants to show that for all but finitely many CM points (x_1, x_2) on C the CM fields of x_1 and x_2 coincide. Then we consider intersections of C with its images under certain Hecke operators. The Riemann hypothesis implies that C is actually contained in some of these images. To finish, we consider an irreducible component X of the inverse image of C in $\mathbb{H} \times \mathbb{H}$, the product of the complex upper half plane by itself, and show that the stabilizer of X in $\mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ is of the kind it should be.

1.4 Remark. At the time this article came back from the referee (June 1997), Yves André succeeded in proving the conclusion of Theorem 1.1 unconditionally, using a result of Masser on Diophantine approximation and the j-function. □

2 Some facts about CM elliptic curves.

Before we start with the proof of Theorem 1.1, we need to recall some facts about CM elliptic curves. These facts can be found for example in [10, Appendix C, §11]. First of all, CM elliptic curves are defined over $\overline{\mathbb{Q}}$. Let K be an imaginary quadratic extension of $\mathbb{Q}$, with a given embedding in $\overline{\mathbb{Q}}$. Let $O_K \subset K$ be the ring of integers. Every subring of O_K of finite index is of the form $O_{K,f} := \mathbb{Z} + fO_K$ for a unique integer $f \geq 1$. For $f \geq 1$ let $S_{K,f}$ be the set of isomorphism classes of pairs (E, α), with E an elliptic curve over $\overline{\mathbb{Q}}$ and $\alpha\colon O_{K,f} \to \mathrm{End}(E)$ an isomorphism of rings inducing the given embedding of K into $\overline{\mathbb{Q}}$ via the action on $\mathrm{Lie}(E)$. The group $G_K := \mathrm{Gal}(\overline{\mathbb{Q}}/K)$ acts on $S_{K,f}$. But also the Picard group $\mathrm{Pic}(O_{K,f})$ acts on $S_{K,f}$ by the following formula:

$$(E, [L]) \mapsto E \otimes_{O_{K,f}} L, \tag{2.1}$$

where L is an invertible $O_{K,f}$-module, $[L]$ its equivalence class and $E \otimes_{O_{K,f}} L$ the cokernel of the map $p\colon E^2 \to E^2$ if $p\colon O_{K,f}^2 \to O_{K,f}^2$ has cokernel L (view p as a matrix with coefficients in $O_{K,f}$). If we choose an embedding of $\overline{\mathbb{Q}}$ in $\mathbb{C}$ and write $E(\mathbb{C})$ as $\mathbb{C}$ modulo a lattice Λ, then $(E \otimes_{O_{K,f}} L)(\mathbb{C})$ is the quotient of $\mathbb{C} \otimes_{O_{K,f}} L$ by $\Lambda \otimes_{O_{K,f}} L$. The actions by G_K and $\mathrm{Pic}(O_{K,f})$ on $S_{K,f}$ commute.

2.2 Proposition. *The set $S_{K,f}$ is a $\mathrm{Pic}(O_{K,f})$-torsor, i.e., the action of $\mathrm{Pic}(O_{K,f})$ is free and has exactly one orbit.*

Proof. (Sketch.) For every (E,α) and Λ as above, $\mathrm{End}_{O_{K,f}}(\Lambda) = O_{K,f}$. Moreover, $O_{K,f}$ is of the form $\mathbb{Z}[x]/(g)$. It follows that Λ is an invertible $O_{K,f}$-module. $\square$

It follows that G_K acts on $S_{K,f}$ via a morphism $G_K \to \mathrm{Pic}(O_{K,f})$. This morphism is surjective and unramified outside f. The Frobenius element at a maximal ideal m not containing f is the element $[m]^{-1}$ of $\mathrm{Pic}(O_{K,f})$ (all this can be seen from deformation theory, using the theorem of Serre-Tate, or from class field theory). Let $H_{K,f}$ be the Galois extension of K corresponding to this quotient $\mathrm{Pic}(O_{K,f})$ of G_K. We remark that we have $H_{K,f} = K(j(E))$ for all (E,α) in $S_{K,f}$.

3 The two CM fields are almost always equal.

Let $C_{\mathbb{C}} \subset \mathbb{C}^2$ be as in Theorem 1.1 (i.e., it is irreducible, it contains infinitely many CM points and its two projections to $\mathbb{C}$ are not constant). Since all CM points have coordinates in $\overline{\mathbb{Q}}$, $C_{\mathbb{C}}$ is defined over $\overline{\mathbb{Q}}$, in the sense that it is the locus of zeros of an irreducible polynomial, call it f, with coefficients in $\overline{\mathbb{Q}}$. It will be convenient for us to work with a curve defined over $\mathbb{Q}$, hence we let C be the union of the finitely many conjugates of $C_{\mathbb{C}}$. Then C is defined by the product F of the Galois conjugates of f, if we take f such that it has a non-zero coefficient in $\mathbb{Q}$. Let d_1 and d_2 be the degrees of F with respect to the second and first variable. Then d_i is the degree of the ith projection from C to $\mathbb{C}$. For x in $\mathbb{C}$ we will denote the endomorphism ring of the corresponding elliptic curve by $\mathrm{End}(x)$. For a CM point x in $\mathbb{C}$ we will call $\mathbb{Q} \otimes \mathrm{End}(x)$ the CM field of x. Note that the isogeny class of a CM elliptic curve over $\overline{\mathbb{Q}}$ consists of all elliptic curves with the same CM field. We want to prove that C is the image in $\mathbb{C}^2$ of some $Y_0(n)$. Our first step in this direction is the following proposition.

3.1 Proposition. *Let C be as above. For all but finitely many CM points (x_1, x_2) in C the CM fields of x_1 and x_2 coincide.*

Proof. Suppose that (x_1, x_2) is a CM point in $C(\overline{\mathbb{Q}})$ such that the two CM fields K_1 and K_2 are different. Since C is defined over $\mathbb{Q}$, $\mathbb{Q}(x_1, x_2)$ has degree at most d_2 over $\mathbb{Q}(x_1)$ and degree at most d_1 over $\mathbb{Q}(x_2)$. Let L be the field generated by K_1 and K_2, and M the intersection of $L(x_1)$ and $L(x_2)$. Let us write $\mathrm{End}(x_i) = O_{K_i,f_i}$ for $i = 1$ and 2. The field $L(x_i)$ is an abelian Galois extension of L, of degree at least $|\mathrm{Pic}(O_{K_i,f_i})|/2$. The degrees of $L(x_1, x_2)$ over $L(x_2)$ and $L(x_1)$ are equal to those of $L(x_1)$ and $L(x_2)$ over M, respectively. This gives us:

$$|\mathrm{Pic}(O_{K_i,f_i})| \leq 2d_i[M : L]. \tag{3.2}$$

We will now work to get a suitable upper bound for $[M : L]$. The group $\mathrm{Gal}(L(x_1, x_2)/\mathbb{Q})$ is an extension of $\mathrm{Gal}(L/\mathbb{Q})$ by the abelian group $\mathrm{Gal}(L(x_1, x_2)/L)$. Hence the action of $\mathrm{Gal}(L(x_1, x_2)/\mathbb{Q})$ on $\mathrm{Gal}(L(x_1, x_2)/L)$ by conjugation factors through an action of $\mathrm{Gal}(L/\mathbb{Q})$.

In the same way, $\mathrm{Gal}(L/\mathbb{Q})$ acts on the two groups $\mathrm{Gal}(L(x_i)/L)$, which we view as subgroups of $\mathrm{Gal}(K_i(x_i)/K_i)$. Now $\mathrm{Gal}(L/\mathbb{Q})$ is equal to $\mathrm{Gal}(K_1/\mathbb{Q}) \times \mathrm{Gal}(K_2/\mathbb{Q})$, hence equal to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. The action of $\mathrm{Gal}(L/\mathbb{Q})$ on $\mathrm{Gal}(L(x_i)/L)$ factors through $\mathrm{Gal}(K_i/\mathbb{Q})$ and as such coincides with the restriction of the action of $\mathrm{Gal}(K_i/\mathbb{Q})$ on $\mathrm{Gal}(K_i(x_i)/K_i) = \mathrm{Pic}(O_{K_i,f_i})$.

3.3 Lemma. *Let K be a quadratic imaginary field and $f \geq 1$. Then the non-trivial element σ of $\mathrm{Gal}(K/\mathbb{Q})$ acts as -1 on $\mathrm{Pic}(O_{K,f})$.*

Proof. The endomorphism $\sigma + 1$ of $\mathrm{Pic}(O_{K,f})$ factors through the norm map from $\mathrm{Pic}(O_{K,f})$ to $\mathrm{Pic}(\mathbb{Z})$. $\square$

Now note that $\mathrm{Gal}(M/L)$ is a quotient of both $\mathrm{Gal}(L(x_i)/L)$, so the action of $\mathrm{Gal}(L/\mathbb{Q})$ on it is by the non-trivial character given by the first projection, but also by the second projection. This implies that $\mathrm{Gal}(M/L)$ is killed by multiplication by two.

3.4 Lemma. *Let K be an imaginary quadratic field and $f \geq 1$. Then the dimension of the $\mathbb{F}_2$-vector space $\mathrm{Pic}(O_{K,f}) \otimes \mathbb{F}_2$ is at most the number of odd primes dividing the discriminant $\mathrm{discr}(O_{K,f})$ of $O_{K,f}$ plus ten.*

Proof. (Sketch.) The exact bound we give does not matter so much, so we just give some indications. First one notes that there is an exact sequence:

$$(K \otimes \mathbb{Q}_2)^* \to \mathrm{Pic}(O_{K,f}) \to \mathrm{Pic}(O_{K,f}[1/2]) \to 0. \tag{3.4.1}$$

Let $S := \mathrm{Spec}(O_{K,f}[1/2])$ and $T := \mathrm{Spec}(\mathbb{Z}[1/2])$. The Kummer sequence gives a surjection from $\mathrm{H}^1(S_{\mathrm{et}}, \mathbb{F}_2)$ onto the 2-torsion subgroup of $\mathrm{Pic}(S)$, which has the same dimension as $\mathrm{Pic}(S) \otimes \mathbb{F}_2$. One deals with $\mathrm{H}^1(S_{\mathrm{et}}, \mathbb{F}_2)$ by projecting to T_{et}. $\square$

Since $\mathrm{Gal}(M/L)$ is killed by 2 and a quotient of a subgroup of $\mathrm{Pic}(O_{K_i,f_i})$, we have:

$$\log_2[M:L] \leq |\{2 \neq p|\mathrm{discr}(O_{K_i,f_i})\}| + 10, \qquad i \in \{1,2\}. \tag{3.5}$$

On the other hand, we have Siegel's theorem (see [8]), stating that:

$$\log|\mathrm{Pic}(O_{K_i,f_i})| = (1/2 + \mathrm{o}(1)) \log|\mathrm{discr}(O_{K_i,f_i})|, \qquad (|\mathrm{discr}(O_{K_i,f_i})| \to \infty). \tag{3.6}$$

Combining equations (3.5) and (3.6) shows that $|\mathrm{Pic}(O_{K_i,f_i})|/[M:L]$ tends to infinity as the discrminiant of O_{K_i,f_i} tends to infinity. But then equation (3.2) can hold for only finitely many (x_1, x_2). This ends the proof of Proposition 3.1. $\square$

3.7 Remark. The proof of Proposition 3.1 shows actually more: the function on the set of CM points on C that sends (x_1, x_2) to f_1/f_2 takes only finitely many values. Using this, one can reduce the proof of Theorem 1.1 to the case where there are infinitely many CM points (x_1, x_2) on C with $\mathrm{End}(x_1) = \mathrm{End}(x_2)$ (one replaces C by its image under a suitable Hecke correspondence). As we do not know how to exploit this, we do not go into further detail. $\square$

3.8 Remark. Proposition 3.1 was also proved by Yves André in [1], and also by Ching-Li Chai (not published). $\square$

4 Intersecting C with something.

We continue the proof of Theorem 1.1. So we let C be as before. At this point we already know that we have infinitely many CM points (x_1, x_2) on C for which x_1 and x_2 are isogeneous because they have the same CM field. We have to prove that there is an integer $n \geq 1$ such that for infinitely many (x_1, x_2) there exists an isogeny of degree n between x_1 and x_2. A direct approach for this is the following. Consider a CM point (x_1, x_2) such that x_1 and x_2 have the same CM field, say K, and an isogeny from x_1 to x_2 of minimal degree, say n. One can get an upper bound for n in terms of the discriminants of the $\mathrm{End}(x_i)$. By Remark 3.7, one can assume that $\mathrm{End}(x_1) = \mathrm{End}(x_2) = O_{K,f}$ and get an upper bound for n from Minkowski's theorem on ideals of small norm representing elements of the class group; the bound is a constant times $|\mathrm{discr}(O_{K,f})|^{1/2}$. Then one considers the intersection of C with $Y_0(n)$. The degrees of both projections from $Y_0(n)$ to $\mathbb{C}$ are equal to $\psi(n)$, where $\psi(n) = n\prod_{p|n}(1+1/p)$. The Picard group of $\mathbb{P}^1 \times \mathbb{P}^1$ (over a field, say $\mathbb{Q}$) is isomorphic to $\mathbb{Z} \times \mathbb{Z}$, the isomorphism sending an effective divisor to the degrees of its two projections to $\mathbb{P}^1$. The intersection form is the following: $(a, b) \cdot (c, d) = ad + bc$. Hence the intersection number of the Zariski closures in $\mathbb{P}^1 \times \mathbb{P}^1$ of C and $Y_0(n)$ is $\psi(n)(d_1 + d_2)$. Since both curves we intersect are defined over $\mathbb{Q}$, the intersection contains all Galois conjugates of (x_1, x_2), of which there are $|\mathrm{Pic}(O_{K,f})|$. So if $|\mathrm{Pic}(O_{K,f})|$ exceeds $\psi(n)(d_1 + d_2)$, the proof is finished, since then the intersection is not proper. Unfortunately, equation (3.6) does not imply such an inequality.

Nevertheless, the idea of intersecting C with something is a good one. Natural "somethings" to take are images of C itself under Hecke correspondences. Again, we consider a CM point (x_1, x_2) on C such that the CM fields of x_1 and x_2 coincide. Let K, f_1 and f_2 be defined by: $\mathrm{End}(x_i) = O_{K,f_i}$. Let f be the least common multiple of f_1 and f_2. The field generated by H_{K,f_1} and H_{K,f_2} is contained in $H_{K,f}$, and one easily checks that $H_{K,f}$ has degree at most three over it. Hence the orbit of (x_1, x_2) under the action of G_K has at least $|\mathrm{Gal}(H_{K,f}/K)|/3$ elements. Recall from §2 that we can identify $\mathrm{Gal}(H_{K,f}/K)$ with $\mathrm{Pic}(O_{K,f})$. For σ in $\mathrm{Gal}(H_{K,f}/K)$ corresponding to the class $[I]$ of an invertible ideal I of $O_{K,f}$, there are isogenies from x_1 to $\sigma(x_1)$ and from x_2 to $\sigma(x_2)$ whose kernels are isomorphic, as $O_{K,f}$-modules, to $O_{K,f}/I$. Hence if we take I such that $O_{K,f}/I$ is a cyclic group of some order n, then $\sigma(x_i)$ is in $T_n(x_i)$ for i equals 1 and 2, where T_n is the correspondence on $\mathbb{C}$ that sends an elliptic curve to the sum (as divisors) of its quotients by its cyclic subgroups of order n. (Let us note that this T_n is not the same as the correspondence on $\mathbb{C}$ that is usually called T_n if n is not square free, since the usual one involves a sum over all subgroups of order n.) Let $T_n \times T_n$ be the correspondence on $\mathbb{C} \times \mathbb{C}$ that is the product of T_n on each factor: it sends a pair (E_1, E_2) of elliptic curves to the sum of the $(E_1/G_1, E_2/G_2)$, where G_i is a cyclic subgroup of order n in E_i. Then (x_1, x_2) is in the intersection of C and $(T_n \times T_n)C$, because x_i is in $T_n(\sigma(x_i))$ and $(\sigma(x_1), \sigma(x_2))$ is in C. Since both C and $(T_n \times T_n)C$ are defined over $\mathbb{Q}$, their intersection contains all Galois conjugates of (x_1, x_2). Hence the intersection has at least $|\mathrm{Pic}(O_{K,f})|/3$ elements. Let us now calculate the degrees of the projections of $(T_n \times T_n)C$ to $\mathbb{C}$. By definition, $(T_n \times T_n)C$ consists of the (x, y) such that there exist u and v in $\mathbb{C}$ with (u, v) in C, and cyclic isogenies of degree n from u to x and from v to y. Let x be in $\mathbb{C}$. Then there are $\psi(n)$ u's with $x \in T_n(u)$. For each such a u there are d_1 v's with (u,v) on C. For each such a v there are $\psi(n)$ y's in $T_n(v)$. This shows that the degree of the first projection of $(T_n \times T_n)C$ is $\psi(n)^2 d_1$. Of course, for the second projection one has the analogous result. So, for the intersection number of C and $(T_n \times T_n)C$ we find $2d_1d_2\psi(n)^2$. We conclude that if $|\mathrm{Pic}(O_{K,f})|$ is bigger than $6d_1d_2\psi(n)^2$, then C is contained in $(T_n \times T_n)C$. The next thing to do is to see if there do exist ideals I with the required properties.

Let x_1, x_2, K and f be as above. Let p be a prime number that splits in $O_{K,f}$, i.e., such that $O_{K,f} \otimes \mathbb{F}_p$ is isomorphic to $\mathbb{F}_p \times \mathbb{F}_p$. For I we take one of the two maximal ideals containing p. As explained above, we have the following implication:

$$6d_1d_2(p+1)^2 < |\mathrm{Pic}(O_{K,f})| \quad \text{implies} \quad C \subset (T_p \times T_p)C. \tag{4.1}$$

Equation (3.6) tells us that $|\mathrm{Pic}(O_{K,f})| = |\mathrm{discr}(O_{K,f})|^{1/2+\mathrm{o}(1)}$. So we want p to be at most something as $|\mathrm{discr}(O_{K,f})|^{1/4}$. More precisely:

4.2 Proposition. *Suppose that there exists $\varepsilon > 0$ such that, when K ranges through all imaginary quadratic fields and f through all positive integers, the number of primes p less than $|\mathrm{discr}(O_{K,f})|^{1/4-\varepsilon}$ that are split in $O_{K,f}$ tends to infinity as $|\mathrm{discr}(O_{K,f})|$ tends to infinity. Then there are infinitely many primes p such that C is contained in $(T_p \times T_p)C$.*

Proof. Because we have infinitely many CM points (x_1, x_2) on C, we know that the discriminants $|\mathrm{discr}(O_{K,f})|$ associated to them as above tend to infinity. The implication (4.1) and equation (3.6) give us the infinitely many required primes. □

5 Existence of small split primes.

The aim of this section is to prove the hypothesis in Proposition 4.2. It turns out that this is no problem at all if one assumes GRH for imaginary quadratic fields and uses the resulting effective Chebotarev theorem of Lagarias, Montgomery and Odlyzko as stated in [9].

For K an imaginary quadratic field, f a positive integer and $x \geq 2$ a real number, let $\pi_{K,f}(x)$ be the number of primes $p \leq x$ that are split in $O_{K,f}$, let $d_K := |\mathrm{discr}(O_K)|$ and let $d_{K,f} := |\mathrm{discr}(O_{K,f})|$. Note that $d_{K,f} = f^2 d_K$. As usual, let $\mathrm{Li}(x) := \int_2^x dt/\log(t)$. Theorem 4 of [9] and the second remark following it say that, for x sufficiently big and for all K as above for which GRH holds, one has:

$$\left|\pi_{K,1}(x) - \frac{1}{2}\mathrm{Li}(x)\right| \leq \frac{1}{6}x^{1/2}\left(\log(d_K) + 2\log(x)\right). \tag{5.1}$$

Since the number of primes dividing f is at most $\log_2(f)$, equation (5.1) implies:

(5.2)
$$\pi_{K,f}(x) \geq \frac{x}{2\log(x)}\left(\mathrm{Li}(x)\frac{\log(x)}{x} - \frac{\log(x)}{3x^{1/2}}\left(\log(d_K) + 2\log(x)\right) - \frac{2\log(x)\log(f)}{x\log(2)}\right).$$

If x tends to infinity, $\mathrm{Li}(x)\log(x)/x$ tends to 1 and $\log(x)^2/x^{1/2}$ tends to 0. One checks easily that for x sufficiently big (i.e., bigger than some absolute constant), and bigger than $\log(d_{K,f})^2(\log(\log(d_{K,f}))^2$, one has $\log(x)\log(d_K)/3x^{1/2} < c < 1$, with c independent of K and f. Under the same conditions, $\log(x)\log(f)/x$ tends to zero if x tends to infinity. This means that we have proved the following proposition.

5.3 Proposition. *Let C be as before (i.e., as in the beginning of §3). Assume GRH for all imaginary quadratic fields. Then there exist infinitely many primes p such that C is contained in $(T_p \times T_p)C$.* □

5.4 Remark. Of course, the question remains whether one can prove the hypothesis of Proposition 4.2 without assuming GRH. Etienne Fouvry tells me the following. He shows that for $r > 0$ and all n, the set of $d_{K,f}$ such that the number of primes $p < d_{K,f}^r$ that are split in $O_{K,f}$ is at most n, has density zero (i.e., the number of such $d_{K,f} < x$ is $\mathrm{o}(x)$ for $x \to \infty$). Moreover, he says that the exponent $1/4$ is critical, in the sense that one can prove that for all $\varepsilon > 0$, the number of primes $p < d_{K,f}^{1/4+\varepsilon}$ that are split in $O_{K,f}$ tends to infinity as $d_{K,f}$ tends to infinity. To prove this, he uses a result of Linnik and Vinogradov in [6], see also [4]. The central point in [6] is an upper bound for short character sums by Burgess, in which the exponent $1/4 + \varepsilon$ appears. This $1/4$ has not moved in the last 30 years. □

6 Some topological arguments.

In this section we finish the proof of Theorem 1.1 by combining Proposition 5.3 with the following theorem, which gives yet another characterization of modular curves.

6.1 Theorem. *Let C in $\mathbb{C}^2$ be an irreducible algebraic curve. Let d_1 and d_2 be the degrees of its two projections to $\mathbb{C}$. Suppose that d_1 and d_2 are both non-zero, and that we have $C \subset (T_n \times T_n)C$ for some square free integer $n > 1$ that is composed of primes $p \geq \max\{5, d_1\}$. Then C is the image of $Y_0(m)$ in $\mathbb{C}^2$ for some $m \geq 1$.*

Let us first show that this theorem and Proposition 5.3 imply Theorem 1.1. So let $C_{\mathbb{C}}$ and C be as in the beginning of §3. Recall that C is the union of the finitely many Galois conjugates of the irreducible component $C_{\mathbb{C}}$ of it. We know that there are infinitely many primes p such that C is contained in $(T_p \times T_p)C$. For such a prime p, let $T_{C,p}$ denote the correspondence on C induced by $T_p \times T_p$. By this we mean the following. The correspondence $T_p \times T_p$ on $\mathbb{C}^2$ is given by the map from $Y_0(p) \times Y_0(p)$ to $\mathbb{C}^2 \times \mathbb{C}^2$ that sends a point (ϕ, ψ) to $(s(\phi), s(\psi), t(\phi), t(\psi))$, where s and t stand for source and target, respectively. Take the inverse image of $C \times C$ in $Y_0(p) \times Y_0(p)$, and delete its zero-dimensional part; that, together with its two maps to C, is $T_{C,p}$. We have to show that a suitable product $T_{C,p_1} \cdots T_{C,p_r}$ with $r \geq 1$ and the p_i distinct induces a non-trivial correspondence from $C_{\mathbb{C}}$ to itself, because then we can apply Theorem 6.1 to $C_{\mathbb{C}}$ with $n = p_1 \cdots p_r$. Let S be the finite set of irreducible components of C. Then each $T_{C,p}$ induces a correspondence $T_{S,p}$ on S that is surjective in the sense that both maps from $T_{S,p}$ to S

are surjective. Moreover, the Galois group $G_{\mathbb{Q}}$ acts transitively on S, and all $T_{S,p}$ are compatible with this action. Let x_0 in S correspond to $C_{\mathbb{C}}$. If there is some $T_{S,p}$ such that x_0 is in $T_{S,p}x_0$, we can take $n = p$. So suppose that for all $T_{S,p}$ we have $x_0 \notin T_{S,p}x_0$. Then we have for all $T_{S,p}$ and all x that $x \notin T_{S,p}x$. One now easily sees that there are $p_1, \ldots, p_r$ distinct with $1 \leq r \leq |S|$ and $x_0 \in T_{S,p_1} \cdots T_{S,p_r}x_0$.

Proof. (Of Theorem 6.1.) We take an integer n as in the theorem we are proving. Let $T_{C,n}$ be the correspondence on C induced by $T_n \times T_n$, in the sense explained above. (In fact, for everything that follows we could also replace $T_{C,n}$ by one of its irreducible components, but it is useful to see how to exploit all of it.) We view $T_{C,n}$ as a subset of $C \times C$. The image of $T_{C,n}$ under the map $(\mathrm{pr}_1, \mathrm{pr}_1)$ from $C \times C$ to $\mathbb{C} \times \mathbb{C}$ is the image T_n of $Y_0(n)$ in $\mathbb{C} \times \mathbb{C}$. Consider the commutative diagram:

$$\begin{array}{ccc} C & \to & \mathbb{C} \\ \uparrow & & \uparrow \\ T_{C,n} & \to & T_n \end{array} \tag{6.2}$$

in which the vertical maps are induced by the projections from $C \times C$ and $\mathbb{C} \times \mathbb{C}$ on the first factor.

6.3 Lemma. *The map from $T_{C,n}$ to the fibred product $C \times_{\mathbb{C}} T_n$ induced by (6.2) is surjective.*

Proof. By construction, all four maps in (6.2) are finite as morphisms of (possibly reducible) algebraic curves. Therefore, the map from $T_{C,n}$ to $C \times_{\mathbb{C}} T_n$ is also a finite morphism of algebraic curves. Hence to show that it is surjective, it suffices to show that $C \times_{\mathbb{C}} T_n$ is irreducible, or, equivalently, that the tensor product of the function fields of C and $Y_0(n)$ over $\mathbb{C}(j)$ is a field. For this, it is enough to prove that the tensor product with $Y_0(n)$ replaced by $Y(n)$ is a field ($Y(n)$ is the modular curve parametrizing elliptic curves with a symplectic basis of their n-torsion). The function field of $Y(n)$ is Galois over $\mathbb{C}(j)$ with Galois group $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})/\{\pm 1\}$. The group $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$ is isomorphic to the product of the $\mathrm{SL}_2(\mathbb{F}_{p_i})$, $1 \leq i \leq r$; one checks easily that it has no non-trivial subgroup of index at most d_1. This means that the function fields of C and $Y(n)$ are linearly disjoint. □

For reasons to become clear soon, we now first prove the following lemma.

6.4 Lemma. *The orbits in C of $T_{C,n}$ are not discrete for the strong topology.*

Proof. The morphism pr_1 from C to $\mathbb{C}$ is proper, hence the image of a closed subset of C is closed in $\mathbb{C}$. In particular, the image of the closure of any subset of C is the closure of its image. Hence it is enough to see that the images in $\mathbb{C}$ of the orbits of $T_{C,n}$ are not closed. Let x be in C, and let y be its image in $\mathbb{C}$. Lemma 6.3 implies that $\mathrm{pr}_1 T_{C,n}x = T_n y$, hence we just have to show that the orbits in $\mathbb{C}$ of T_n are not closed. For this we view $\mathbb{C}$ as the quotient of the complex upper half plane $\mathbb{H}$ by the group $\mathrm{SL}_2(\mathbb{Z})$ via the map $\pi\colon \tau \mapsto j(\mathbb{C}/(\mathbb{Z} + \mathbb{Z}\tau))$. Let x be in $\mathbb{C}$, and choose τ in $\pi^{-1}x$. Then for all a and b in $\mathbb{Z}$, $\pi(\tau+a)$ and $\pi(n^b\tau)$ are in the orbit of x under T_n. By composing these operations, we see that $\pi(n^b\tau+a)$ and $\pi(\tau+n^{-b}a)$ are in the orbit of x. Taking a non-zero and b big shows that the orbit is not closed. (In fact, it is easy to show, using $\tau\mapsto -\tau^{-1}$, that all orbits in $\mathbb{C}$ of T_n are dense.) $\square$

We view $\mathbb{C}\times\mathbb{C}$ as the quotient of $\mathbb{H}\times\mathbb{H}$ by the group $\Gamma := \mathrm{SL}_2(\mathbb{Z})\times\mathrm{SL}_2(\mathbb{Z})$, via the map:

$$\pi\colon \mathbb{H}\times\mathbb{H}\to\mathbb{C}\times\mathbb{C},\quad (\tau_1,\tau_2)\mapsto (j(\mathbb{C}/(\mathbb{Z}+\mathbb{Z}\tau_1)), j(\mathbb{C}/(\mathbb{Z}+\mathbb{Z}\tau_2))). \tag{6.5}$$

Let X be an irreducible component of the analytic subvariety $\pi^{-1}C$ of $\mathbb{H}\times\mathbb{H}$. The group $G := \mathrm{SL}_2(\mathbb{R})\times\mathrm{SL}_2(\mathbb{R})$ acts transitively on $\mathbb{H}\times\mathbb{H}$. We will study its subgroup G_X, the stabilizer of X. What we have to prove is that G_X is the graph of an inner automorphism of $\mathrm{SL}_2(\mathbb{R})$; this automorphism then tells us for which m our curve C is the image of $Y_0(m)$. The decisive step in the proof of this is to see that G_X is not discrete (if C is an arbitrary curve in $\mathbb{C}^2$, then G_X is typically discrete).

6.6 Lemma. *The group G_X is an analytic subgroup of G.*

Proof. The action of G on $\mathbb{H}\times\mathbb{H}$ is algebraic (it is given by fractional linear transformations). The subgroup G_X consists of exactly those elements g in G that satisfy, for all x in X, the two conditions $gx\in X$ and $g^{-1}x\in X$. All these conditions are analytic. $\square$

6.7 Lemma. *The kernels of the two projections from G_X to $\mathrm{SL}_2(\mathbb{R})$ are discrete.*

Proof. This kernel K, say for the second projection, is the same as the stabilizer of X in the subgroup $\mathrm{SL}_2(\mathbb{R})\times\{1\}$ of G. For all τ in $\mathbb{H}$, it stabilizes $X_\tau := X\cap(\mathbb{H}\times\{\tau\})$, which is discrete since $d_2>0$; hence the connected component K^o of K stabilizes every element of X_τ. We conclude that K^o acts trivially on X. Now the stabilizer in $\mathrm{SL}_2(\mathbb{R})$ of the element i of $\mathbb{H}$ is $\mathrm{SO}_2(\mathbb{R})$. Because $d_1>0$, K^o is contained in all conjugates of $\mathrm{SO}_2(\mathbb{R})$, the intersection of which is $\{\pm 1\}$. $\square$

6.8 Lemma. *The image in $\mathrm{SL}_2(\mathbb{Z})$ of Γ_X, the stabilizer of X in Γ, under the ith projection, has index at most d_i.*

Proof. We do the proof for $i=2$. We factor the map $\pi\colon\mathbb{H}\times\mathbb{H}\to\mathbb{C}\times\mathbb{C}$ as follows:

$$\mathbb{H}\times\mathbb{H}\to\mathbb{C}\times\mathbb{H}\to\mathbb{C}\times\mathbb{C}. \tag{6.8.1}$$

Let Y be the image of X in $\mathbb{C}\times\mathbb{H}$. Then Y is an irreducible component of the inverse image Z of C in $\mathbb{C}\times\mathbb{H}$. Let S be the set of c in C such that every x in $\pi^{-1}c$ is contained in more than one irreducible component of $\pi^{-1}C$. Then S is contained in the finite subset of C consisting of singular points and points of which at least one of the coordinates is in $\{0,1728\}$. Let C' be $C-S$, and let X' and Y' be the inverse images, in X and Y, respectively, of C'. The map from X' to C' is the quotient for the action of Γ_X, hence the map from Y' to C' is the quotient for the action of $\mathrm{pr}_2\Gamma_X$. It follows that $\mathrm{pr}_2\Gamma_X$ is the stabilizer in $\mathrm{SL}_2(\mathbb{Z})$ of Y in Z, so the set $\mathrm{SL}_2(\mathbb{Z})/\mathrm{pr}_2\Gamma_X$ is the set of irreducible components of Z. But Z is also the fibred product of $\mathrm{pr}_2\colon C\to\mathbb{C}$ and $\mathbb{H}\to\mathbb{C}$, which implies that Z has at most d_2 irreducible components. $\square$

Lemmas 6.6, 6.7 and 6.8 are in fact valid for any curve C in $\mathbb{C}^2$ for which d_1 and d_2 are non-zero. The next one crucially exploits that $C \subset (T_n \times T_n)C$.

6.9 Lemma. *The topological group G_X is not discrete.*

Proof. The subgroup G_X of G is analytic, hence closed. It contains Γ_X. The inclusion $C \subset (T_n \times T_n)C$ implies that it contains some less trivial elements as well. The correspondence T_n on $\mathbb{C}$ can be described as follows. Take z in $\mathbb{C}$; take its inverse image in $\mathbb{H}$; apply the map $\tau \mapsto n\tau = \left(\begin{smallmatrix} n & 0 \\ 0 & 1 \end{smallmatrix}\right)\tau$ to it and take its image in $\mathbb{C}$; that is $T_n z$. Another way to say this is: take representatives t_i in $\mathrm{GL}_2(\mathbb{Q})$ (there are $\psi(n)$ of them) for the quotient set $\mathrm{SL}_2(\mathbb{Z})\left(\begin{smallmatrix} n & 0 \\ 0 & 1 \end{smallmatrix}\right)\mathrm{SL}_2(\mathbb{Z})/\mathrm{SL}_2(\mathbb{Z})$; then for z in $\mathbb{C}$ and τ in $\mathbb{H}$ mapping to it, $T_n z$ is the image of the sum of the $t_i\tau$. It follows that for each (i, j) such that $(t_i, t_j)X$ is contained in $\pi^{-1}C$ we get an element $g_{i,j}$ in G_X of the form

$$g_{i,j} = \gamma_{i,j,1} \cdot \left(n^{-1/2}\begin{pmatrix} n & 0 \\ 0 & 1 \end{pmatrix}, n^{-1/2}\begin{pmatrix} n & 0 \\ 0 & 1 \end{pmatrix}\right) \cdot \gamma_{i,j,2},$$

with $\gamma_{i,j,1}$ and $\gamma_{i,j,2}$ in Γ. For c in C and x in X mapping to c, $T_{C,n}c$ is the image of the sum of the $g_{i,j}x$. Let H be the subgroup of G_X generated by Γ_X and these elements $g_{i,j}$. We will prove that H is not discrete. Let $\overline{H}$ be the closure of H. We take an element x in X. The map from G to $\mathbb{H} \times \mathbb{H}$ sending g to gx is proper, because the stabilizers of elements of $\mathbb{H} \times \mathbb{H}$ are compact. Hence $\overline{H}x$ is also the closure of Hx. The subset Hx of X is discrete if and only if its image in C is discrete, since H contains Γ_X and the map $X \to C$ is the quotient for the action of Γ_X. By construction, the image of Hx in C is the orbit of x for $T_{C,n}$, which, by Lemma 6.4, is not discrete. This proves that G_X is not discrete. $\square$

We can now quickly finish the proof of Theorem 6.1. Consider the Lie algebra $\mathrm{Lie}(G_X)$, which by Lemma 6.9 is non-zero. Lemma 6.7 tells us that the two projections $\mathrm{pr}_i\mathrm{Lie}(G_X)$ are non-zero. But $\mathrm{pr}_i\mathrm{Lie}(G_X)$ is normalized by $\mathrm{pr}_i\Gamma_X$, which is Zariski dense in $\mathrm{SL}_2(\mathbb{R})$ by Lemma 6.8. Since $\mathrm{Lie}(\mathrm{SL}_2(\mathbb{R}))$ is simple, it follows that $\mathrm{pr}_i\mathrm{Lie}(G_X)$ is equal to $\mathrm{Lie}(\mathrm{SL}_2(\mathbb{R}))$ for both i. So, since $\mathrm{SL}_2(\mathbb{R})$ is connected, G_X projects surjectively on both factors $\mathrm{SL}_2(\mathbb{R})$ of G. Now we apply what is called Goursat's lemma: let H be a subgroup of a product $G_1 \times G_2$, such that the projections p_1 and p_2 from H to G_1 and G_2 are surjective, then $\ker(p_1)$ and $\ker(p_2)$ are normal subgroups of G_2 and G_1, respectively, and H is the inverse image of the graph of an isomorphism between $G_1/\ker(p_2)$ and $G_2/\ker(p_1)$. The kernel of $\mathrm{pr}_2\colon G_X \to \mathrm{SL}_2(\mathbb{R})$ is a normal subgroup of $\mathrm{SL}_2(\mathbb{R})$, viewed as $\mathrm{SL}_2(\mathbb{R}) \times \{1\}$. Since it is discrete and contains $\{1, -1\}$, it is $\{1, -1\}$. The same holds for the other projection, and G_X is the inverse image in G of the graph of an analytic automorphism, σ say, of $\mathrm{SL}_2(\mathbb{R})/\{\pm 1\}$. Every such automorphism is inner. Since the $\mathrm{pr}_i\Gamma_X$ have finite index in $\mathrm{SL}_2(\mathbb{Z})$, it follows that σ is induced from an inner automorphism of the algebraic group $\mathrm{SL}_{2,\mathbb{Q}}$. The algebraic group of automorphisms of $\mathrm{SL}_{2,\mathbb{Q}}$ is $\mathrm{PGL}_{2,\mathbb{Q}}$. Since the map $\mathrm{GL}_2(\mathbb{Q}) \to \mathrm{PGL}_2(\mathbb{Q})$ is surjective (for example by Hilbert 90), σ is given by conjugation by some element g in $\mathrm{GL}_2(\mathbb{Q})$. So G_X is the set $\{(h, \pm ghg^{-1}) \mid h \in \mathrm{SL}_2(\mathbb{R})\}$. Let x be an element of X, and write it as $x = (\tau, h\tau)$ with τ in $\mathbb{H}$ and h in $\mathrm{SL}_2(\mathbb{R})$.

Since $G_X x$ is in X, which is of dimension two, the stabilizer of x in G_X has dimension at least one. Let H be the stabilizer of τ in the connected component of identity G_X^o, for the action of G_X^o on the first factor $\mathbb{H}$; then the stabilizer of $h\tau$ for the action on the second factor is the conjugate $g^{-1}hHh^{-1}g$ of H. Since H is of dimension one and connected (it is isomorphic to $\mathrm{SO}_2(\mathbb{R})$) we must have $H = g^{-1}hHh^{-1}g$, i.e., $g^{-1}h$ normalizes H. Since the normalizer of $\mathrm{SO}_2(\mathbb{R})$ in $\mathrm{SL}_2(\mathbb{R})$ is just $\mathrm{SO}_2(\mathbb{R})$ itself, this means that $g^{-1}h$ is in H, or, equivalently, that $h\tau = g\tau$. This means that $X = \{(\tau, g\tau) \mid \tau \in \mathbb{H}\}$. We may replace g by multiples ag of it, with a a non-zero rational number. So we can and do suppose that $g\mathbb{Z}^2$ is contained in $\mathbb{Z}^2$ and that $\mathbb{Z}^2/g\mathbb{Z}^2$ is cyclic, say of order m. It is now clear that C is $Y_0(m)$. □

7 Some remarks.

7.1 Remark. Our proof of Theorem 1.1 shows in fact that, assuming GRH, for each pair (d_1, d_2) of positive integers there exists an effectively computable number $B(d_1, d_2)$, such that on every irreducible curve C in $\mathbb{C}^2$ of bi-degree (d_1, d_2) that is defined over $\mathbb{Q}$ and not a modular curve there are at most $B(d_1, d_2)$ CM points. (Note that under GRH, the statement that $|\mathrm{Pic}(O_K)|/|\mathrm{Pic}(O_K)[2]| \to \infty$ is effective.) □

7.2 Remark. It is not true that all irreducible curves C in $\mathbb{C}^2$ with $C \subset (T_n \times T_n)C$ for some $n > 1$ are the image of some $Y_0(m)$. Here we construct some examples. Let $n > 1$. Let w_n be the Atkin-Lehner involution of $Y_0(n)$: it sends an isogeny to its dual. The correspondence T_n on $\mathbb{C}$ has the following description. For z in $\mathbb{C}$, take its inverse image in $Y_0(n)$, take the image of that under w_n and then the image in $\mathbb{C}$. It follows that for an irreducible curve C in $\mathbb{C}^2$ such that at least one of the irreducible components of its inverse image in $Y_0(n) \times Y_0(n)$ is stable under the involution (w_n, w_n) we have $C \subset (T_n \times T_n)C$. Let Z be the quotient of $Y_0(n) \times Y_0(n)$ by that involution. Bertini's theorem, see for example [5, Theorem 6.3], gives the existence of whole families of curves in Z with irreducible inverse image in $Y_0(n) \times Y_0(n)$. Take C to be the image in $\mathbb{C}^2$ of such an inverse image. □

7.3 Remark. The condition that n be square free in Theorem 6.1 should not be necessary; it is due to the laziness of the author. □

7.4 Remark. It is very tempting to try to generalize the methods of this article to the general case of Oort's conjecture. □

Acknowledgements. I would like to thank Rutger Noot for interesting discussions on this subject that motivated me enough to work on it, and for his remarks on previous versions of this article. I thank Johan de Jong for his interest and the reference to [9]. I am very grateful to Etienne Fouvry for a letter in which he explains in detail the results mentioned in Remark 5.4. Tim Dokshitzer pointed out a gap in a previous version of this article. I want to thank Fabrice Rouillier for helping me installing the necessary software on the computer with which this article is written. Finally, I am very grateful for an invitation to the Centre for Research in Mathematics at the Institut d'Estudis Catalans in Barcelona, where I could compile my somewhat chaotic and incomplete notes into this article.

References

[1] Y. André. Talk given at the IHP in Paris on January 28, 1996.

[2] Y. André. *G-functions and geometry.* Aspects of mathematics, Vol. **E13**, Vieweg, Braunschweig, 1989.

[3] J.W.S. Cassels and A. Frohlich, eds. *Algebraic number theory.* Academic Press, 1967.

[4] J. Friedlander. *Certain hypotheses concerning L-functions.* Pacific Journal of Mathematics **69** (1977), 37–44

[5] J.P. Jouanolou. *Théorèmes de Bertini et applications.* Progress in Mathematics, Vol. **42**, Birkhauser, Boston, 1983.

[6] Yu.V. Linnik and A.I. Vinogradov. *Hyperelliptic curves and the least prime quadratic residue.* Doklady Nauk. Akad. SSSR **168** (1966), 259–261.

[7] B.J.J. Moonen. *Special points and linearity properties of Shimura varieties.* Thesis, Utrecht, September 1995.

[8] J. Oesterlé. *Nombres de classes des corps quadratiques imaginaires.* Séminaire Bourbaki, exposé 631, juin 1984. Astérisque **121–122** (1985), 309–323.

[9] J-P. Serre. *Quelques applications du théorème de densité de Chebotarev.* Publications Mathématiques de l'IHES **54** (1981), 123–202.

[10] J.H. Silverman. *The arithmetic of elliptic curves.* Graduate Texts in Mathematics **106**, Springer-Verlag, 1986.

Bas Edixhoven
IRMAR
Campus de Beaulieu
35042 Rennes cedex
France