Mathematical Research Letters **10**, 145–150 (2003)

MODULARITY OF ABELIAN SURFACES WITH QUATERNIONIC MULTIPLICATION

Luis Dieulefait

Abstract. We prove that any abelian surface defined over $\mathbb{Q}$ of GL_2-type having quaternionic multiplication and good reduction at 3 is modular. We generalize the result to higher dimensional abelian varieties with "sufficiently many endomorphisms."

1. Statement of the theorem

In this brief article we interest ourselves in the relation to classical modular forms of the following geometric object: A an abelian surface defined over $\mathbb{Q}$ such that $\mathrm{End}^0_{\mathbb{Q}}(A) = \mathbb{Q}(\sqrt{d})$ (d a non-square integer) and $\mathrm{End}^0_{\overline{\mathbb{Q}}}(A) = B$ where $B/\mathbb{Q}$ is an indefinite rational quaternion algebra.

We borrow from P. Clark's note (see [C]) the name "premodular" QM-surfaces over $\mathbb{Q}$ for these abelian surfaces. The condition of having real multiplication defined over $\mathbb{Q}$ is necessary (and sufficient) in order to obtain a two dimensional Galois representation of the full $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on the Tate modules of A. This condition has to be imposed because there are examples of QM-abelian surfaces defined over $\mathbb{Q}$ that are not premodular (and a fortiori they are not modular, see [DR2]). Observe that the action of the quaternion algebra can not be defined over $\mathbb{Q}$ because $\mathbb{Q} \subseteq \mathbb{R}$. In fact (we keep from now on the premodularity condition) the minimal field K such that $\mathrm{End}^0_K(A) = B$ is an imaginary quadratic field (cf. [DR1]).

A generalization of the Shimura-Taniyama conjecture predicts that any abelian variety of GL_2-type over $\mathbb{Q}$ is modular (see [R2]), thus as a particular case modularity of all premodular QM-surfaces over $\mathbb{Q}$ is expected.

Examples of this kind of surfaces have been constructed and studied by Brumer (see [HHM]), Hashimoto and Murabayashi (see [HM]) and Bending.

The fact that K is imaginary implies that the field $\mathbb{Q}(\sqrt{d})$ has to be real (cf. [R2], Prop. (7.2), second proof). This result combined with a description of rationality fields of endomorphisms of QM-abelian surfaces (see [DR1]) tells us that a principally polarized premodular QM-surface over $\mathbb{Q}$ can only exist if the quaternion algebra B is "exceptional", i.e., there exists $m \mid D = \mathrm{disc}(B)$ such that $B = (\frac{-D,m}{\mathbb{Q}})$.

In this article we will prove the following result:

Received July 3, 2002.

Theorem 1.1. *Any premodular QM-surface over $\mathbb{Q}$ having good reduction at 3 is modular.*

Remark. J. Ellenberg has obtained, as a consequence of his proof of Serre's conjecture over $\mathbb{F}_9$ under local conditions at 3 and 5, a much stronger result asserting the modularity of every abelian surface over $\mathbb{Q}$ having real multiplication whenever the surface has good ordinary reduction at 3 and 5 (see [E]). Our result is not contained in his because it improves these local conditions, but it is much more restrictive because of the QM-multiplication hypothesis.

Another result of modularity of premodular QM-surfaces was also obtained by Hasegawa, Hashimoto and Momose (see [HHM]) with the condition that it exists an odd prime ramifying in the quaternion algebra B such that the surface has good reduction at it. Thanks to this result, we can assume in the proof of our theorem that B is unramified at 3.

2. The proof

Let A be a premodular QM-abelian surface over $\mathbb{Q}$ with $\mathrm{End}^0_{\mathbb{Q}}(A) = \mathbb{Q}(\sqrt{d})$, $d > 0$, $\mathcal{O}$ the ring of integers of $\mathbb{Q}(\sqrt{d})$ and B the indefinite quaternion algebra inside $\mathrm{End}^0_{\overline{\mathbb{Q}}}(A)$. Let N be the product of the primes of bad reduction of A and $D = \mathrm{disc}(B)$. From now on we will assume $3 \nmid N \cdot D$.

In the following lines, we shall recollect a few facts about Galois representations attached to A (see [R1], [R2], [HHM], [Die] and [DR1] for references).

For any rational prime ℓ and $\lambda \mid \ell$ a prime in $\mathcal{O}$ the Galois-action on the ℓ-adic Tate module of A (because $\mathrm{End}^0_{\mathbb{Q}}(A) = \mathbb{Q}(\sqrt{d})$) gives a representation

$$\rho_\lambda : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathcal{O}_\lambda)$$

which is odd, irreducible, and unramified outside $\ell \cdot N$. The field $\mathbb{Q}(\sqrt{d})$ being real, we have: $\det \rho_\lambda = \chi$, the ℓ-adic cyclotomic character (cf. [R2]).

Further restrictions on the compatible family of Galois representations $\{\rho_\lambda\}$ are imposed by the fact that A has QM. The situation is like in the case of a modular form having a single inner twist (studied by Momose, Ribet, Papier): the field of definition of the QM-action is a quadratic field K, more precisely, an imaginary quadratic field unramified outside N. Let H be the Galois group with fixed field K. For every $\ell \nmid N \cdot D$, the image of $\rho_\lambda|_H$ is a normal subgroup of index at most 2 of the full image of ρ_λ. On the other hand, for every $\ell \nmid N \cdot D$

$$\rho_\lambda|_H : \mathrm{Gal}(\overline{\mathbb{Q}}/K) \to \mathrm{GL}_2(\mathbb{Z}_\ell)$$

and this restriction is surjective for almost every ℓ (cf. [Oh] and [DR1]). If we call (for every $p \nmid N$) a_p the trace of the image of $\rho_\lambda(\mathrm{Frob}\ p)$, $\lambda \nmid pND$, we know that $\mathbb{Q}(\{a_p\}) = \mathbb{Q}(\sqrt{d})$ and the following condition is satisfied:

$$a_p^\gamma = \varphi(p) a_p \tag{2.1}$$

for almost every $p \nmid N$, where γ is the order two element in $\mathrm{Gal}(\mathbb{Q}(\sqrt{d})/\mathbb{Q})$ and φ is the quadratic character corresponding to $K/\mathbb{Q}$. Serre proved that this is equivalent to $\mathbb{Q}(\{a_p^2\}) = \mathbb{Q} \subsetneq \mathbb{Q}(\{a_p\})$.

From this, it follows (cf. [R1] and [DR1]) that for $\ell \nmid N \cdot D$, $\lambda \mid \ell$, the image of ρ_λ is contained in the subgroup of $\mathrm{GL}_2(\mathcal{O}_\lambda)$ generated by $\mathrm{GL}_2(\mathbb{Z}_\ell)$ and the diagonal matrix:

$$\begin{pmatrix} a_p & 0 \\ 0 & a_p^{-1} \end{pmatrix}$$

for some $p \nmid \ell N$ verifying $\lambda \nmid a_p$, $a_p \neq 0$ and $\varphi(p) = -1$ (if such a prime p exists). By (2.1), these two last conditions are equivalent to: $a_p = u_p\sqrt{d}$, $u_p \in \mathbb{Z}[\frac{1}{2}]$, $u_p \neq 0$.

Let $\bar{\rho}_\lambda$ be the reduction modulo λ of ρ_λ (take it semisimple). If for all primes $p \nmid \ell N$ with $\varphi(p) = -1$ we have $\lambda \mid a_p$, instead of the description above we will make use of the fact that (after suitable conjugation) the image of $\bar{\rho}_\lambda$ will be contained in $\mathrm{GL}_2(\mathbb{F}_\ell)$. This follows from the fact that in this case formula (2.1) implies that all traces are in $\mathbb{F}_\ell$, and the representation being odd (and $\det(\bar{\rho}_\lambda) = \chi$) it is well known that this implies that it has a model over $\mathbb{F}_\ell$ (cf. for example [W1], [W2] or [R3]). Observe finally that for a residual representation the property of being modular depends only on the traces, so it is stable under conjugation.

Thanks to the above description, we see that (as in the proof of modularity of $\mathbb{Q}$-curves by Ellenberg and Skinner, see [ES]) we are in a comfortable situation because we have "for free" that the residual representation $\bar{\rho}_t$, if absolutely irreducible, is modular, where t is a prime in $\mathcal{O}$ dividing 3. This follows from the theorem of Langlands and Tunnell as in Wiles' original proof (see [W2] and also [ES]), because even if in general the image of $\bar{\rho}_t$ is not contained in $\mathrm{GL}_2(\mathbb{F}_3)$ its projectivisation $\mathbb{P}(\bar{\rho}_t)$ has image in $\mathrm{PGL}_2(\mathbb{F}_3)$.

We divide the rest of the proof in two cases:

1) $\bar{\rho}_t$ absolutely irreducible:

As we already explained, we know that the residual representation $\bar{\rho}_t$ is modular. If we also assume that the restriction to $\mathbb{Q}(\sqrt{-3})$ remains absolutely irreducible, then all conditions are satisfied to apply Diamond's generalization of Taylor-Wiles modularity result (see [D], [W2], [TW]) and conclude that ρ_t is modular, and therefore that A is modular.

Thus, we assume that the restriction of $\bar{\rho}_t$ to $\mathbb{Q}(\sqrt{-3})$ is absolutely reducible. This implies that we are in the "dihedral case", namely, that the image of $\bar{\rho}_t$ is contained in the normaliser $\mathcal{N}$ of a Cartan subgroup C of $\mathrm{GL}_2(\bar{\mathbb{F}}_3)$, but not contained in C. We also know that the restriction to $\mathbb{Q}(\sqrt{-3})$ of our representation has its image inside C.

Thus, the composition of $\bar{\rho}_t$ with the quotient $\mathcal{N}/C$:

$$\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathcal{N} \twoheadrightarrow \mathcal{N}/C \tag{2.2}$$

gives the quadratic character corresponding to $\mathbb{Q}(\sqrt{-3})$. In particular, this quadratic character ramifies at 3.

We know that, in general, for any $\ell \nmid N$, the restriction of the residual representation $\bar{\rho}_\lambda$ to the inertia subgroup at ℓ has only two possibilities (this is a classical result due to Raynaud with some restrictions on λ, which follows in general for $\ell > 2$ from the fact that the λ-adic representation is Barsotti-Tate, therefore crystalline with Hodge-Tate weights 0 and 1, via an application of Fontaine-Laffaille theory, cf. [FL] and [B], chapter 9):

$$\bar{\rho}_\lambda|_{I_\ell} \cong \begin{pmatrix} \chi & * \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \psi_2 & 0 \\ 0 & \psi_2^\ell \end{pmatrix}$$

where ψ_2 is a fundamental character of level 2.

If we suppose that $\bar{\rho}_t|_{I_3}$ acts through level 2 fundamental characters, the image of I_3 gives a cyclic group of order $4 > 2$, thus (this is a standard trick of Serre and Ribet, see [R3]) it has to be contained in C. But this implies that the quadratic character defined by composition (2.2) should be unramified at 3, contradicting the fact that this character corresponds to $\mathbb{Q}(\sqrt{-3})$.

Thus, we can assume that we are in the first case of Raynaud's result: $\bar{\rho}_t|I_3$ is reducible. Then, we use again the fact that the representations ρ_λ are crystalline with Hodge-Tate weights 0 and 1 for every $\ell \nmid N$, $\ell > 2$, to conclude, applying results of Breuil (see [B], chapter 9), that the local representation $\rho_t|_{G_3}$ is reducible (we can apply this result of classification of crystalline representations at $p = 3$ because the highest Hodge-Tate weight is $w = 1$ and so: $3 > w + 1$). Therefore, all conditions are satisfied to apply the result of Skinner and Wiles (see [SW2]) on nearly ordinary deformations of (residual) modular irreducible Galois representations to conclude that ρ_t, thus A, is modular.

2) $\bar{\rho}_t$ absolutely reducible:

In this case, we exclude again the second possibility in Raynaud' s theorem. This time this is automatic (Kronecker-Weber) and we have:

$$\bar{\rho}_t^{s.s.} = \epsilon \oplus \epsilon^{-1}\chi$$

where χ is the mod 3 cyclotomic character and ϵ a character ramifying only at primes in N.

Thus, excluded the case of level 2 fundamental characters, we apply again the results of Breuil (cf. [B], chapter 9) on crystalline representations to conclude that $\rho_t|_{G_3}$ is reducible. Now, an application of the result of Skinner and Wiles (see [SW1]) on deformations of residually reducible Galois representations proves that ρ_t, thus A, is modular. □

3. An example

The following genus two curve (taken from a family of curves constructed by Brumer) has the property that its Jacobian has real multiplication by $\mathbb{Q}(\sqrt{5})$ defined over $\mathbb{Q}$:

$$y^2 = x^6 + 47x^5 + 365x^4 + 865x^3 + 400x^2 + 38x - 4$$

The primes of bad reduction are $2, 5$ and 127. Let us call A its Jacobian. This example was considered in [HHM], were it is proved that A verifies $\mathrm{End}^0_{\overline{\mathbb{Q}}}(A) = \mathrm{End}^0_{\mathbb{Q}(\sqrt{-10})}(A) = B$ where B is the indefinite rational quaternion algebra of discriminant 10. Computing a few characteristic polynomials we see that $\mathrm{End}_{\mathbb{Q}}(A) = \mathbb{Z}[\sqrt{5}]$. Applying theorem 1.1 we conclude that A is modular.

Remark. The results of [HHM] can not be applied to prove the modularity of A because the primes ramifying in B are also primes of bad reduction. The result of [E] is also insufficient because A has bad reduction at 5.

4. Generalization

For higher dimensional abelian varieties of GL_2-type having large enough endomorphism algebras, we can still use exactly the same argument to prove modularity. For example, consider the case of an abelian variety A of dimension 2^n defined over $\mathbb{Q}$ with:

(I) $\mathrm{End}^0_{\mathbb{Q}}(A)$ a totally real number field E of degree 2^n such that:

(II) $\mathrm{End}^0_{\overline{\mathbb{Q}}}(A) = M_{2^{n-1}}(B)$, with $B/\mathbb{Q}$ an indefinite quaternion algebra.

Following [R2], we know that $\mathrm{Gal}(E/\mathbb{Q})$ is an abelian group of exponent 2 and that the traces $a_p := \mathrm{trace}(\rho_\lambda(\mathrm{Frob}\ p))$ again verify: $\mathbb{Q}(\{a_p^2\}) = \mathbb{Q}$ and the two dimensional Galois representations $\{\rho_\lambda\}$ have, for every $\lambda \nmid D = \mathrm{disc}(B)$, the following property: the image of the projectivization of the residual representation $\bar{\rho}_\lambda$ is contained in $\mathrm{PGL}_2(\mathbb{F}_\ell)$. Therefore, from the argument given in section 2 we conclude:

Theorem 4.1. *Let A be an abelian variety of dimension 2^n verifying* (I) *and* (II) *above. Assume that A has good reduction at* 3*. Then A is modular.*

Remark. Again, we can assume in the proof that $3 \nmid D = \mathrm{disc}(B)$ (if $3 \mid D$ modularity follows from [HHM], Theorem 2.1).

Final Remark. One of the referees suggested that the arguments and results in this article can be easily extended to the case of multiplicative reduction at 3. Instead of applying the results of Breuil, all the desired statements about restrictions of ρ_λ to decomposition groups should follow in this case from the Mumford-Tate uniformization.

References

[B] C. Breuil, *p-adic Hodge theory, deformations and local Langlands,* notes of a course at CRM, Bellaterra, Spain, July 2001.

[C] P. Clark, *Two Quaternionic modular philosophies,* preprint: `http://modular.fas.harvard.edu/mcs/archive/Fall2001/notes/10-15-01/`

[D] F. Diamond, *On deformation rings and Hecke rings,* Ann. of Math. (2) **144** (1996), 137–166.

[Die] L. Dieulefait, *Newforms, inner twists, and the inverse Galois problem for projective linear groups,* J. Théor. Nombres Bordeaux **13** (2001), 395–411.

[DR1] L. Dieulefait, V. Rotger, *The arithmetic of QM-abelian surfaces through their Galois representations,* preprint.

[DR2] ———, *Non-modular fake elliptic curves,* preprint.

[E] J. Ellenberg, *Serre's conjecture over* $\mathbb{F}_9$*,* preprint: `front.math.ucdavis.edu/math.NT/0107147`

[ES] J. Ellenberg, C. Skinner, *On the modularity of Q-curves,* Duke Math. J. **109** (2001), 97–122.

[FL] J. M. Fontaine, G. Laffaille, *Construction de représentations p-adiques,* Ann. Sci. École Norm. Sup. (4) **15** (1982), 547–608 (1983).

[HHM] Y. Hasegawa, K. Hashimoto, F. Momose, *Modularity conjecture for Q-curves and QM-curves,* Internat. J. Math. **10** (1999), 1011–1036.

[HM] K. Hashimoto, N. Murabayashi, *Shimura curves as intersections of Humbert surfaces and defining equations of QM-curves of genus two,* Tohoku Math. J. (2) **47** (1995), 271–296.

[Oh] M. Ohta, *On ℓ-adic representations of Galois groups obtained from certain two dimensional abelian varieties,* J. Fac. Sci. Univ. Tokyo Sect. IA Math. **21** (1974), 299–308.

[R1] K. Ribet, *On ℓ-adic representations attached to modular forms. II,* Glasgow Math. J. **27** (1985), 185–194.

[R2] ———, *Abelian varieties over* $\mathbb{Q}$ *and modular forms,* Algebra and topology 1992 (Taejŏn), 53–79, Korea Adv. Inst. Sci. Tech., Taejŏn, 1992.

[R3] ———, *Images of semistable Galois representations,* Pacific J. Math. **1997**, Special Issue, 277–297.

[SW1] C. Skinner, A. Wiles, *Residually reducible representations and modular forms,* Inst. Hautes Études Sci. Publ. Math. **89** (1999), 5–126 (2000).

[SW2] ———, *Nearly ordinary deformations of irreducible residual representations,* Ann. Fac. Sci. Toulouse Math. (6) **10** (2001), 185–215.

[TW] R. Taylor, A. Wiles, *Ring-theoretical properties of certain Hecke algebras,* Ann. of Math. (2) **141** (1995), 553–572.

[W1] A. Wiles, *On ordinary λ-adic representations associated to modular forms,* Invent. Math. **94** (1988), 529–573.

[W2] ———, *Modular elliptic curves and Fermat's last theorem,* Ann. of Math. (2) **141** (1995), 443–551.

Institut Galilée, Université de Paris 13, 99, av. J.B. Clément, 93430 Villetaneuse, France.

E-mail address: `luisd@math.univ-paris13.fr`

E-mail address: `luisd@mat.ub.es`