Modularity of fibres in rigid local systems

Henri Darmon

September 5, 2007

1 Introduction

Let K be a totally real field embedded in a fixed algebraic closure $\overline{K}$, and write $G_K := \mathrm{Gal}(\overline{K}/K)$ for its absolute Galois group. Fix a prime $\ell \neq 2$, and consider an odd two-dimensional Galois representation

$$\rho : G_K \longrightarrow \mathbf{GL}_2(E),$$

where E is either a finite field of characteristic ℓ or a finite extension of $\mathbb{Q}_\ell$. Assume that the restrictions of ρ to the inertia groups at the primes of K above ℓ are *potentially semistable* in the sense of [FM].

The representation ρ is called *modular* if it is associated to a Hilbert modular form on $\mathbf{GL}_2(K)$, as is explained, for example, in [W1] and [W2]. Fontaine and Mazur [FM] conjectured that this is always the case. Significant progress on this conjecture was achieved [W3] by proving particular instances of the following "lifting conjecture":

Conjecture 1.1 *Suppose that ℓ is odd and that the residual representation $\bar{\rho}$ attached to ρ is modular. Then ρ itself is modular.*

Conjecture 1.1 is proved in [W3] and [TW] when $K = \mathbb{Q}$ and the restriction of ρ to the decomposition groups at the primes above ℓ are *semistable* in the sense of [DDT], sec. 2.4. This is enough (using the primes $\ell = 3$ and 5) to establish the Shimura-Taniyama conjecture for semistable elliptic curves, thanks to a key result of Langlands and Tunnell. Progressively stronger cases of conjecture 1.1 were subsequently proved by [Di], [CDT], [Fu], and [SW]; in [SW], Skinner and Wiles obtain quite general results in the context where

K is any totally real field, the principal assumption being that ρ is *ordinary* at the primes above ℓ.

In this note we consider Galois representations which occur in "rigid families", and establish their modularity under conjecture 1.1. This implies the modularity (over suitable real abelian extensions) of the Galois representations occuring in the cohomology of the curves

$$y^n = x^a(x-1)^b(x-t)^c, \qquad t \in \mathbb{Q},$$

whose periods as a function of the parameter t are values of classical hypergeometric functions.

To state the main result precisely, denote by $K(t)$ the field of rational functions in the indeterminate t, and let

$$\varrho : G_{K(t)} \longrightarrow \mathbf{GL}_2(E)$$

be a two-dimensional Galois representation. For $x \in \mathbb{P}^1(\bar{K})$, viewed as a place of $K(t)$, let $D_x \subset G_{K(t)}$ be a decomposition group at x, and write $I_x = \hat{\mathbb{Z}}(1)$ for its inertia subgroup. One says that ϱ is *unramified* at x if its restriction to I_x is trivial. If in addition x belongs to $\mathbb{P}^1(K)$, the restriction of ϱ to D_x factors through $D_x/I_x = G_K$, giving rise to a Galois representation

$$\varrho[x] : G_K \longrightarrow \mathbf{GL}_2(E),$$

which can be thought of as the specialization of ϱ at $t = x$.

Let ϱ^{geom} be the restriction of ϱ to the subgroup

$$G^{geom} := \mathrm{Gal}(\overline{K(t)}/\overline{K}(t)) \subset G_{K(t)}.$$

The representation ϱ is said to be *rigid* if ϱ^{geom} is unramified outside 0, 1, and ∞. (The reason for this terminology will be made clear in the next section, cf. prop. 2.4.) Choose a topological generator of $\hat{\mathbb{Z}}(1)$ corresponding to a compatible system (ζ_n) of primitive n-th roots of unity. For $j = 0, 1, \infty$, let γ_j be the corresponding generator of I_j, and let $\sigma_j = \varrho(\gamma_j) \in \mathbf{GL}_2(E)$. The *monodromy matrices* σ_j depend on the choice of decomposition groups D_j but their conjugacy classes in $\mathbf{GL}_2(E)$ are well-defined. We will show (lemma 2.2) that the semisimplification of σ_j has finite order n_j. One can then prove (prop. 2.4) that the "field of definition" K of ϱ necessarily contains

the real subfield $K_{n_j} := \mathbb{Q}(\zeta_{n_j})^+$ of the cyclotomic field of n_j-th roots of unity. Conversely, ϱ has a twist which extends to a representation of $G_{K_n(t)}$, where $n = n(\varrho)$ is the least common multiple of the n_j. Replace ϱ by such a twist, and K by K_n. Our main result is then:

Theorem 1.2 *Let ϱ be a rigid representation, and assume that one of the σ_j is unipotent, and that 8 does not divide $n = n(\varrho)$.*

If conjecture 1.1 is true, then $\varrho[x]$ arises from a Hilbert modular form over K_n, for all $x \in \mathbb{P}^1(\mathbb{Q}) - \{0, 1, \infty\}$.

Acknowledgements. The author thanks Nick Katz for helpful conversations and the ETH in Zürich for its hospitality while this article was written. This research was funded by grants from NSERC and by an Alfred P. Sloan research fellowship.

2 Rigid representations

Fix a rigid representation ϱ, and keep the notations of the introduction. While the monodromy matrices σ_j are only defined up to conjugation, the decomposition groups D_j can be chosen so that the relation

$$\sigma_0\sigma_1\sigma_\infty = 1 \tag{1}$$

is satisfied (cf. for example [Se1], th. 6.3.2.) Fix such a choice from now on.

A 2×2 matrix is called a *reflection* if its eigenvalues are 1 and -1.

Lemma 2.1 *The matrix σ_j is either a reflection or an element of $\mathbf{SL}_2(E)$.*

Proof: The conjugacy classes of σ_j are *rational* over the real field K in the sense of [Se1], sec. 7.1. In particular, σ_j is conjugate to σ_j^{-1}; the result follows.

If one of the σ_j is a reflection, then exactly two are, because of the relation $\det(\sigma_0\sigma_1\sigma_\infty) = 1$ which follows from (1). In that case the image of ϱ^{geom} is a dihedral group. We exclude this case from consideration from now on, and assume that each σ_j belongs to $\mathbf{SL}_2(E)$. The matrix σ_j is said to be *quasi-unipotent* if its minimal polynomial has a double root.

Lemma 2.2 *The σ_j are either quasi-unipotent or of finite order.*

Proof: Let $K^{cyc} := K(\zeta_\infty)$ be the maximal cyclotomic extension of K, and let Ω be its Galois group, identified with a subgroup of $\hat{\mathbb{Z}}^\times$. Since the conjugacy class of σ_j is rational over K, the matrix σ_j is conjugate to σ_j^α for all $\alpha \in \Omega$. But Ω has finite index in $\hat{\mathbb{Z}}^\times$, and hence the eigenvalues of σ_j are roots of unity.

Definition 2.3 *An* admissible triple *in* $\mathbf{SL}_2(E)$*) is a triple* $(\sigma_0, \sigma_1, \sigma_\infty)$ *of elements in* $\mathbf{SL}_2(E)$*, taken modulo conjugation in* $\mathbf{GL}_2(E)$*, and satisfying*
(a) The semisimplification of σ_j *has finite order* n_j*;*
(b) The group generated by σ_0*,* σ_1*, and* σ_∞ *is an irreducible subgroup of* $\mathbf{SL}_2(E)$*.*
(c) $\sigma_0\sigma_1\sigma_\infty = 1$*.*

Let $n = n(\varrho) = \mathrm{lcm}(n_0, n_1, n_\infty)$, as before. The following "rigidity" property justifies the terminology of the introduction.

Proposition 2.4 *Let* $(\sigma_0, \sigma_1, \sigma_\infty)$ *be an admissible triple in* $\mathbf{SL}_2(E)$ *with* σ_1 *unipotent. Then there exists a rigid representation*

$$\varrho : G_{K_n(t)} \longrightarrow \mathbf{GL}_2(E)$$

whose monodromy matrix at $t = j$ *is equal to* σ_j*. Furthermore, if* ϱ' *is any irreducible rigid representation whose monodromy matrices are conjugate to those of* ϱ*, then* ϱ' *is conjugate to* $\varrho \otimes \chi$*, where* $\chi : G_{K_n} \longrightarrow E^\times$ *is a constant central character.*

This follows from theorems 1 and 2 of [Be]. (See also the discussion in section 1 of [Da2].)

3 Hypergeometric abelian varieties

For the following definition, let K be any real abelian field, and $\mathcal{O}_K$ its ring of integers. (We will also write $\mathcal{O}_n := \mathbb{Z}[\zeta_n + \zeta_n^{-1}]$ to denote the ring of integers of K_n.)

Definition 3.1 *A* hypergeometric abelian variety with multiplications by K *is an abelian scheme* A *over* $(\mathbb{P}^1 - \{0, 1, \infty\})_{/\mathbb{Q}}$ *of dimension* $[K : \mathbb{Q}]$ *equipped with an inclusion*

$$\iota : \mathcal{O}_K \hookrightarrow End_{K(t)}(A)$$

which is compatible with the natural action of $Gal(K/\mathbb{Q})$ on both sides, and whose associated monodromy representation is irreducible.

Define an *admissible triple* $(\sigma_0, \sigma_1, \sigma_\infty)$ in $\mathbf{SL}_2(\mathcal{O}_K)$ in the obvious way (replacing E by $\mathcal{O}_K$ in definition 2.3). Given a hypergeometric abelian variety A with multiplications by K, one can associate to it an admissible triple $(\sigma_0, \sigma_1, \sigma_\infty)$ in $\mathbf{SL}_2(\mathcal{O}_K)$ by letting σ_j be the image of γ_j acting on the DeRham cohomology $H^1_{Dr}(A)$ (viewed as a two-dimensional K-vector space). Conversely, given an admissible triple $(\sigma_0, \sigma_1, \sigma_\infty)$ in $\mathbf{SL}_2(\mathcal{O}_K)$, let n_j be the order of the semisimplification of σ_j and set $n = \mathrm{lcm}(n_0, n_1, n_\infty)$. One sees that K must contain the fields K_{n_j} generated by the traces of the σ_j. Assume that $K = K_n$.

Proposition 3.2 *Assume that σ_1 is unipotent. There exists a hypergeometric abelian variety A with multiplications by K_n whose associated monodromies are $(\sigma_0, \sigma_1, \sigma_\infty)$. The isogeny class of this abelian variety depends only on the triple $(\sigma_0, \sigma_1, \sigma_\infty)$ (modulo conjugation by $\mathbf{GL}_2(E)$).*

Proof: See [Ka], sec. 5.4, or [CW], sec. 3.3. The hypergeometric abelian varieties are constructed as appropriate quotients of the Jacobians of the curves

$$y^n = x^a(x-1)^b(x-t)^c.$$

From hypergeometric abelian varieties to rigid representations: If A is a hypergeometric abelian variety with multiplications by K, the ℓ-adic Tate module

$$T_\ell(A) := \lim_{\leftarrow} A[\ell^k]$$

is a free module of rank two over $\mathcal{O}_K \otimes \mathbb{Z}_\ell$. The natural action of $G_{\mathbb{Q}(t)}$ on this Tate module is semilinear, in the sense that

$$\alpha(s \cdot v) = s^\alpha \cdot \alpha(v), \quad \text{for} \quad \alpha \in G_{\mathbb{Q}(t)}, \quad s \in \mathcal{O}_K \otimes \mathbb{Z}_\ell, \quad v \in T_\ell(A).$$

In particular, if φ is a homomorphism from $\mathcal{O}_K$ to E, then $T_\ell(A) \otimes_\varphi E$ is a two-dimensional E-vector space on which $G_{K(t)}$ acts linearly. It gives rise to a rigid two-dimensional Galois representation ϱ of $G_{K(t)}$, and thus to a family of representations $\varrho[x]$ of G_K for all $x \in K - \{0, 1\}$.

Definition 3.3 *The hypergeometric abelian variety A is said to be* modular *at x if $\varrho[x]$ is associated to a Hilbert modular form over K with coefficients in E, for all choices of (φ, E). We say that A is* modular *if it is modular at x, for all $x \in \mathbb{Q}$.*

Remark: The representations $\varrho[x]$ attached to A, as ℓ, E, and φ vary, form a *compatible system* of ℓ-adic representations of G_K, and hence to prove that A is modular at x, it suffices to prove that $\varrho[x]$ is modular for a single $E \subset \bar{\mathbb{Q}}_\ell$.

Examples:
1. If σ_0, σ_1, and $\sigma_\infty \in \mathbf{SL}_2(\mathbb{Z})$ are quasi-unipotent with eigenvalues 1, 1, -1, then A is isogenous to the Legendre family of elliptic curves

$$y^2 = x(x-1)(x-t).$$

The modularity of A is thus a special case of the Shimura-Taniyama conjecture which was completely established by Wiles [W3].

2. If σ_0 and σ_∞ are of order 4 and 3 respectively, and σ_1 is unipotent, then $A_{/\mathbb{Q}(t)}$ is isogenous to (a twist of) the universal family of elliptic curves of invariant $j = 1728/(t-1)$. The modularity of A in this case is merely a re-formulation of the Shimura-Taniyama conjecture.

3. If σ_0 and σ_1 are unipotent and σ_∞ is of order r with r an odd prime, the corresponding hypergeometric abelian variety is the Jacobian of the hyperelliptic curve with real multiplications by $\mathbb{Q}(\zeta_r)^+$ given by the equation

$$y^2 = (x+2)(f(x) + 2 - 4t),$$

where $f(x) = xg(x^2-2)$ and $g(x)$ is the characteristic polynomial of $-(\zeta_r + \zeta_r^{-1})$. This curve had already been considered in [TTV], and used in [Da2] to study the generalized Fermat equation $x^p + y^p = z^r$. In the language of [Da2], the mod p representations attached to A are the "even Frey representations" associated to the generalized Fermat equation $x^p + y^p = z^r$.

4. If σ_0 and σ_1 are unipotent and σ_∞ has order $2r$ with r an odd prime, then A is the Jacobian of the hyperelliptic curve (also used in the study of $x^p + y^p = z^r$)

$$y^2 = f(x) + 2 - 4t.$$

From rigid representations to hypergeometric abelian varieties:

Let ϱ be a rigid representation of $G_{K_n}(t)$ with unipotent monodromy at $t = 1$, associated to an admissible triple $(\sigma_0, \sigma_1, \sigma_\infty)$ in $\mathbf{SL}_2(E)$. This triple can be lifted to an admissible triple $(\tilde{\sigma}_0, \tilde{\sigma}_1, \tilde{\sigma}_\infty)$ in $\mathbf{SL}_2(\mathcal{O}_n)$, i.e., there is a homomorphism $\varphi : \mathcal{O}_n \longrightarrow E$ such that $\varphi(\tilde{\sigma}_j) = \sigma_j$, and $\tilde{\sigma}_1$ is unipotent. Let A be the hypergeometric abelian variety with multiplications by K_n associated to $(\tilde{\sigma}_0, \tilde{\sigma}_1, \tilde{\sigma}_\infty)$ by proposition 3.2. Then we have:

Proposition 3.4 *The representation ϱ is equivalent to (a twist of) the Galois representation obtained from the action of $G_{K(t)}$ on $T_\ell(A) \otimes_\varphi E$.*

Proof: This is a direct consequence of the uniqueness statement of proposition 2.4, since the representation associated to $T_\ell(A) \otimes_\varphi E$ is a rigid representation associated to the triple $(\sigma_0, \sigma_1, \sigma_\infty)$.

Thanks to proposition 3.4, it is enough to show that all hypergeometric abelian varieties with unipotent monodromy at $t = 1$ are modular in order to prove theorem 1.2.

4 Congruences

Let A be a hypergeometric abelian variety with multiplication by $K = K_n$, and let $(\sigma_0, \sigma_1, \sigma_\infty)$ be the associated admissible triple in $\mathbf{SL}_2(\mathcal{O}_K)$. Assume that σ_1 is unipotent, and let ℓ be an odd prime which divides $n = \mathrm{lcm}(n_0, n_1, n_\infty)$. For $j = 0, 1, \infty$, let n'_j be the prime-to-ℓ part of n_j, let n' be the prime-to-ℓ part of n, and let $K' = \mathbb{Q}(\zeta_{n'})^+$. Choose a prime λ of K above ℓ, and let λ' be the unique prime of K' below it. The prime λ' is totally ramified in K/K', so that the residue fields of K and K' at λ and λ' respectively are canonically isomorphic. Let $\mathbb{F}$ be this common residue field. It is equipped with maps $\varphi : \mathcal{O}_K \longrightarrow \mathbb{F}$ and $\varphi' : \mathcal{O}_{K'} \longrightarrow \mathbb{F}$. Let $(\sigma'_0, \sigma'_1, \sigma'_\infty)$ be a lift of $(\varphi(\sigma_0), \varphi(\sigma_1), \varphi(\sigma_\infty))$ to an admissible triple in $\mathbf{SL}_2(\mathcal{O}_{K'})$, and let A' be the abelian variety associated to it by proposition 3.2.

Because $G_{\mathbb{Q}(t)}$ acts semi-linearly on $A[\ell] \otimes_\varphi \mathbb{F}$ and because λ is totally ramified in K/K', the action of $G_{K(t)}$ on this $\mathbb{F}$-vector space extends to a linear action of $G_{K'(t)}$.

Theorem 4.1 *The $G_{K'(t)}$ representation $A[\ell] \otimes_\varphi \mathbb{F}$ is isomorphic to (a twist of) the representation $A'[\ell] \otimes_{\varphi'} \mathbb{F}$.*

Proof: A direct consequence of proposition 3.4.

5 Proof of the main result

Theorem 5.1 *Let A be a hypergeometric abelian variety with multiplications by K_n, and let $(\sigma_0, \sigma_1, \sigma_\infty)$ be the associated admissible triple. Assume that σ_1 is unipotent, and that 8 does not divide n. If conjecture 1.1 is true, then A is modular.*

Proof: The proof is by induction on $d = [K_n : \mathbb{Q}]$. If $d = 1$, then A is an elliptic curve over $\mathbb{Q}(t)$ and the modularity of A follows from the Shimura-Taniyama conjecture, which itself follows from conjecture 1.1. If $d > 1$, then n is divisible by an odd prime ℓ, by the assumption that 8 does not divide n. Adopting the notation of section 4, we begin by showing (for a fixed $t = x \in \mathbb{Q}$) that $A[\ell] \otimes_\varphi \mathbb{F}$ is associated to a Hilbert modular form f_ℓ over K. If $A[\ell] \otimes_\varphi \mathbb{F}$ is a reducible representation of G_K, then one may express f_ℓ in terms of Eisenstein series. Assume that $A[\ell] \otimes_\varphi \mathbb{F}$ is irreducible. Since $n' < n$ and $d' = [K' : \mathbb{Q}] < d$, the induction hypothesis implies that A' is modular. Hence so is the rigid representation $A'[\ell] \otimes_{\varphi'} \mathbb{F}$; let f'_ℓ be the associated Hilbert modular form mod ℓ on $\mathbf{GL}_2(K')$. By theorem 4.1, the $G_{K'}$ module $A[\ell] \otimes_\varphi \mathbb{F}$ is isomorphic to $A'[\ell] \otimes_{\varphi'} \mathbb{F}$, and so corresponds to the same f'_ℓ. Letting f_ℓ be the cyclic base change lift (from K' to K) of f'_ℓ, it follows that the representation $A[\ell] \otimes_\varphi \mathbb{F}$ is modular over K. The λ-adic Tate module $T_\ell(A) \otimes K_\lambda$ is a potentially semistable Galois representation, since it arises from the torsion points of an abelian variety. Hence it is modular, by conjecture 1.1.

Remark: The proof that A is modular involves repeated applications of the lifting conjecture 1.1, once with each odd prime ℓ dividing n. In light of the results in [SW], it might be feasible to prove unconditionally the modularity of A at $t = x$, when x is such that A is *ordinary* at all these primes. There are infinitely many values of x with this property: for example, all the x for which n divides the numerator of $x - 1$.

References

[Be] G.V. Belyi, On Galois extensions of a maximal cyclotomic field. *Math. USSR Izvestija*, **14** (1980) No. 2, 247–256.

[CDT] B. Conrad, F. Diamond, R. Taylor, Modularity of certain potentially Barsotti-Tate Galois representations, Journal of the AMS, to appear.

[CW] P. Cohen, J. Wolfart, Modular embeddings for some non-arithmetic Fuchsian groups, *Acta Arithmetica* **LVI** (1990) pp. 93-110.

[Da2] H. Darmon, Rigid local systems, Hilbert modular forms, and Fermat's Last Theorem, CICMA preprint.

[DDT] H. Darmon, F. Diamond, and R. Taylor, Fermat's Last Theorem, *Current Developments in Mathematics* **1**, 1995, International Press, pp. 1-157.

[Di] F. Diamond, On deformation rings and Hecke rings. *Annals of Math* (2) **144** (1996), no. 1, 137–166.

[FM] J.-M. Fontaine, B. Mazur, Geometric Galois representations, in *Elliptic curves, modular forms, and Fermat's Last theorem*, International Press, Cambridge, 1995, 41–78.

[Fu] K. Fujiwara, Deformation rings and Hecke algebras in the totally real case, preprint.

[Ka] N. Katz, Exponential sums and differential equations, Annals of Mathematics Studies, No. 124, Princeton, NJ, 1990.

[Se1] J.-P. Serre, Topics in Galois Theory. Jones and Bartlett, 1992.

[SW] C. Skinner, A. Wiles, *Ordinary representations and modular forms.* Proc. Nat. Acad. Sci. U.S.A. 94 (1997), no. 20, 10520–10527.

[TTV] W. Tautz, J. Top, A. Verberkmoes, Explicit hyperelliptic curves with real multiplication and permutation polynomials. *Canad. J. Math.* **43** (1991) no. 5, 1055-1064.

[TW] R. Taylor and A. Wiles, Ring-theoretic properties of certain Hecke algebras. *Ann. of Math.* (2) **141** (1995), no. 3, 553–572.

[W1] A. Wiles, On ordinary λ-adic representations associated to modular forms. *Invent. Math.* **94** (1988), no. 3, 529–573.

[W2] A. Wiles, On p-adic representations for totally real fields. *Ann. of Math.* (2) **123** (1986), no. 3, 407–456.

[W3] A. Wiles, Modular elliptic curves and Fermat's last theorem. *Ann. of Math.* (2) **141** (1995), no. 3, 443–551.