*Journal de Théorie des Nombres*
*de Bordeaux* **15** (2003), 367–381

# Modularity of Galois representations

par Chris SKINNER

Résumé. Dans cet article, nous donnons une interprétation en termes de théorie de Galois des représentations galoisiennes $p$-adiques de dimension 2 associés aux formes modulaires holomorphes de Hilbert qui sont des "new forms". L'article suit pour l'essentiel l'exposé des Journées Arithmétiques de 2001.

Abstract. This paper is essentially the text of the author's lecture at the 2001 Journées Arithmétiques. It addresses the problem of identifying in Galois-theoretic terms those two-dimensional, $p$-adic Galois representations associated to holomorphic Hilbert modular newforms.

By the work of many mathematicians (Eichler, Shimura, Deligne, Carayol ... ) it is known that if $f$ is a holomorphic Hilbert modular newform over a totally real field $F$ of degree $d$ and if the weight $(k_1, \ldots, k_d)$ of $f$ is such that the $k_i$'s are greater than zero and all have the same parity, then for each rational prime $p$ there is a continuous representation $\rho_{f,p} : \mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ such that trace $(\mathrm{frob}_\ell) = c_\ell(f)$ for almost all primes $\ell$ of $F$, where $c_\ell(f)$ is the eigenvalue of the action on $f$ of the usual Hecke operator $T_\ell$. Conversely, it is expected that any "reasonable" representation $\rho : \mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ should be isomorphic to some $\rho_{f,p}$ (see the Conjecture in §1 below). The first significant result of this type was obtained by A. Wiles [W1] for the case $F = \mathbf{Q}$. This paper discusses efforts and results extending those in Wiles' ground-breaking work. Emphasis is placed on loosening hypotheses (e.g., replacing $\mathbf{Q}$ with an arbitrary $F$ or allowing reducible residual representations).

## 1. Introduction

This lecture focuses on modularity in the context of Galois representations. More to the point, it discusses various recently-proven results pertaining to the problem of intrinsically characterizing the $p$-adic Galois representations associated to automorphic representations.

Manuscrit reçu le 30 janvier 2002.

We begin by recalling some of the properties of the Galois representations associated to holomorphic Hilbert modular newforms. Suppose that $F$ is a totally real number field of degree $d$ and that $f$ is a holomorphic Hilbert modular newform of level $\mathfrak{n}$, Nebentypus $\chi$, and weight $(k_1, ..., k_d)$ with each $k_i \geq 1$ and all having the same parity. For each prime ideal $\ell$ of $F$ let $c_\ell(f)$ be the eigenvalue of the usual $\ell^{\text{th}}$ Hecke operator $T_\ell$ acting on $f$ (so $T_\ell f = c_\ell(f)f$). It is known by the work of many mathematicians (Eichler, Shimura, Deligne, Carayol, Wiles, Taylor, Blasius, Rogawski, Tunnell, Jarvis...) that for each rational prime $p$, upon fixing an embedding $\overline{\mathbf{Q}} \hookrightarrow \overline{\mathbf{Q}}_p$ (so we may view the $c_\ell(f)$'s as elements of $\overline{\mathbf{Q}}_p$), there is a continuous representation

$$\rho_{f,p} : \mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$$

such that

(i) $\rho_{f,p}$ is irreducible,
(ii) $\det \rho_{f,p}(c) = -1$ for all complex conjugations $c$ (i.e., $\rho_{f,p}$ is *odd*),
(iii) $\det \rho_{f,p} = \chi\varepsilon^m$, $m = \max\{k_i - 1\}$,
(iv) $\mathrm{trace}\, \rho_{f,p}(\mathrm{frob}_\ell) = c_\ell(f)$, $\ell \nmid \mathfrak{n}p$.

In (iii), we view $\chi$ as a Galois character via class field theory, with the reciprocity map normalized so that arithmetic Frobenii ($\mathrm{frob}_\ell$) correspond to uniformizers. Also, $\varepsilon$ is the usual $p$-adic cyclotomic character.

The results described herein are mostly concerned with the following problem.

**Problem**. *Intrinsically characterize the set $\mathcal{MG}_{F,p}$ of continuous representations $\rho : \mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ such that $\rho \cong \rho_{f,p}$ for some Hilbert modular newform $f$.*

By an *intrinsic characterization* we mean a characterization in terms of the Galois-theoretic properties of the representation (such a characterization is found in the conjecture stated in §1 below).

The celebrated proof of the modularity of elliptic curves defined over $\mathbf{Q}$ (a consequence of the main results of [W1], [TW] and their extensions in [D1], [CDT], and [BCDT]) entailed showing that the representation attached to the Tate module for some prime $p$ (and hence for all primes) for such a curve is contained in $\mathcal{MG}_{\mathbf{Q},p}$. This result has been amply discussed elsewhere and was a topic of a lecture by B. Conrad at the previous Journées Arithmétiques. The organizers of the current conference requested that the author discuss other aspects of efforts to resolve the above problem, so this is the last that will be said about elliptic curves until the final section.

The organization of the rest of this lecture is essentially as follows. Since it is often easier to couch results in terms of automorphic representations (and results in the literature are often expressed in such terms), we first reformulate the properties of $\rho_{f,p}$ in these terms and then go on to state a conjectural resolution of the problem. Then the related notion of a residual representation is introduced. This is done in §1. In §2 we summarize the strategy introduced and successfully implemented by A. Wiles to prove many cases of the conjecture is summarized. In §3 we discuss results of K. Fujiwara and of Wiles and the author in the case where the residual representation is irreducible but $F$ may not be equal to $\mathbf{Q}$. In §4 the work of Wiles and the author in the case where the residual representation is reducible is described. Finally, in §5 we mention a few things that the constraints of space and time (and of the author's understanding!) have forced us to omit from the preceding sections.

## 2. Automorphic representations and Galois representations

To simplify matters, fix once and for all an embedding $\overline{\mathbf{Q}} \hookrightarrow \mathbf{C}$ and compatible embeddings $\overline{\mathbf{Q}} \hookrightarrow \overline{\mathbf{Q}}_\ell$ and $\overline{\mathbf{Q}}_\ell \hookrightarrow \mathbf{C}$ for each prime $\ell$. Let $F$ be a totally real number field of degree $d$ with ring of integers $\mathcal{O}_F$. Suppose that $\pi = \otimes\pi_v$ is a cuspidal automorphic representation of $\mathrm{GL}_{2/F}$ such that

(i) if $v|\infty$ then $\pi_v$ is a discrete series representation with weight $k_v \geq 1$ and central character $\psi_{\pi_v}$ satisfying $\psi_{\pi_v}(-1) = (-1)^{k_v}$,
(ii) the $k_v$'s, $v|\infty$, all have the same parity,
(iii) the central character of $\pi$ is $\chi_\pi|\cdot|_{\mathbf{A}}^{m-1}$ where $\chi_\pi$ is finite and $m = \max\{k_v - 1\}$.

There is a one-to-one correspondence between such representations and the Hilbert modular newforms as in the introduction.

For each prime $p$ there is a continuous representation

$$\rho_{\pi,p} : \mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$$

such that

$$\text{(1.1)} \qquad \begin{array}{l} \text{(i) } \rho_{\pi,p} \text{ is irreducible,} \\ \text{(ii) if } v \nmid p\infty,\ \rho_{\pi,p}|_{W_v} \cong \rho(\pi_v), \end{array}$$

where $W_v$ is the Weil group at the finite place $v$ and $\rho(\pi_v)$ is the representation of $W_v$ associated to $\pi_v$ via the local Langlands correspondence (cf. [Ta, Theorem 4.2.1]), which we normalize so that if $\pi_v = \pi(\chi_1, \chi_2)$ is unramified then $\rho(\pi_v) \cong \chi_1^{-1}|\cdot|_v^{-1/2} \oplus \chi_2^{-1}|\cdot|_v^{-1/2}$. If $\pi$ corresponds to the newform $f$, then $\rho_{\pi,p}$ is just $\rho_{f,p}$, properties (ii), (iii), and (iv) of $\rho_{f,p}$ following from property (ii) of (1.1).

One can also say something about $\rho_{\pi,p}|_{D_v}$ for $v|p$, where $D_v$ is a decomposition group at $v$, namely that in many cases (and conjecturally in all)

(1.2) if $v|p$ then $\rho_{\pi,p}|_{D_v}$ is *potentially semistable.*

The definition of "potentially semistable" is fairly technical (cf. [Fo]) but reflects properties of $p$-adic representations that appear in the cohomology of algebraic varieties. (That $\rho_{\pi,p}$ does when each $k_v \geq 2$ is essentially the content of [BR]).

The problem from the introduction is thus to characterize the continuous representations $\rho : \mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ such that $\rho \cong \rho_{\pi,p}$ for some $\pi$. A guess at a solution can be obtained by extracting as many Galois-theoretic properties as possible from (1.1) and (1.2) and then positing that $\mathcal{MG}_{F,p}$ consists of all $p$-adic representations having these properties. This leads to the following conjecture (which is essentially Conjecture 3c of [FoM]).

**Conjecture**. *Let $\mathcal{MG}^T_{F,p}$ be the set of Tate twists of representations in $\mathcal{MG}_{F,p}$. The set $\mathcal{MG}^T_{F,p}$ consists of the continuous representations $\rho$ : $\mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ such that*

(i) *$\rho$ is irreducible,*
(ii) *$\rho$ is odd,*
(iii) *$\rho$ is unramified at all but finitely many places,*
(iv) *for all $v|p$, $\rho|_{D_v}$ is potentially semistable.*

In order to state various theorems in the direction of the above Conjecture, we need to introduce the notion of a residual representation and discuss some specific instances of potential semistability.

Suppose $\rho : G \to \mathrm{GL}_n(\overline{\mathbf{Q}}_p)$ is a continuous representation of a compact, Hausdorff group. A simple measure-theoretic argument shows that some conjugate of $\rho$ takes values in $\mathrm{GL}_n(\mathcal{O})$ with $\mathcal{O}$ the ring of integers of some finite extension of $\mathbf{Q}_p$ containing the eigenvalues of all elements of $\rho(G)$. Replacing $\rho$ with this conjugate, we define the *residual representation* $\overline{\rho}$ of $\rho$ to be the semisimplification of the representation into $\mathrm{GL}_n(k)$, $k$ the residue field of $\mathcal{O}$, obtained by reducing the matrix entries of $\rho$ modulo the maximal ideal of $\mathcal{O}$. This is well-defined up to equivalence and up to extension of the field $k$.

Suppose $\rho : \mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ is continuous. For the rest of this lecture we will be concerned with two possibilities for $\rho|_{D_v}$ whenever $v|p$. The first will be that

$$\rho|_{D_v} \cong \begin{pmatrix} \chi_1 & * \\ & \chi_2 \end{pmatrix}, \quad \chi_2|_{I_v} \text{has finite order},$$

$I_v$ being an inertial group at $v$. In this instance $\rho$ is said to be *nearly ordinary at $v$*. We will say that such a $\rho$ is *$D_v$-distinguished* if $\overline{\chi}_1 \neq \overline{\chi}_2$. From work of Wiles [W2] (see also [H2]) it follows that if $\pi = \otimes \pi_v$ is such that the $k_v$, $v|\infty$, are all equal and that for all $v|p$, $\pi_v \cong \pi(\xi_1, \xi_2)$ is either a principal series representation or a special representation with $\xi_1/\xi_2 = |\cdot|_v$

satisfying $\xi_2(\lambda_v)|\lambda_v|_v^{1/2}$ is a unit in the ring of integers of $\overline{\mathbf{Q}}_p$ ($\lambda_v$ a fixed uniformizer of $\mathcal{O}_{F,v}$), then $\rho_{\pi,p}$ is nearly ordinary at all $v|p$.

The second is that

$\rho|_{D_v}$ is the extension of scalars of a representation
arising from a $p$-divisible group over $\mathcal{O}_{F,v}$.

In this instance $\rho$ is said to be *flat at* $v$. Note that $\rho$ can be both nearly ordinary and flat at $v$. From work of Carayol and Taylor (see [T]), it follows that if $\pi = \otimes\pi_v$ is such that $k_v = 2$ for all $v|\infty$ and $\pi_v$ is unramified for all $v|p$, then $\rho_{\pi,p}$ is flat at all $v|p$ if either the degree of $F$ is odd or if $\overline{\rho}_{\pi,p}$ is irreducible.

I can now state the first significant results in the direction of the Conjecture. These were obtained by Wiles [W1], [TW] and Taylor and subsequently refined in [D1] to yield

**Theorem 1** ([W1],[D1]). *Suppose $p$ is an odd prime. If $\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ is a continuous representation such that*

(i) *$\rho$ is irreducible and unramified at all but finitely many primes,*
(ii) *$\rho$ is odd,*
(iii) *$\det\rho = \chi\varepsilon^m$ with $\chi$ finite and $m \geq 1$,*
(iv) *$\rho$ is either flat at $p$ or nearly ordinary at $p$ and $D_p$-distinguished,*
(v) *$\overline{\rho}|_{\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}(\sqrt{\left(\frac{-1}{p}\right)p}))}$ is irreducible,*
(vi) *$\overline{\rho}\cong\overline{\rho}_{\pi,p}$ for some $\pi$,*

*then $\rho \in \mathcal{MG}_{\mathbf{Q},p}$.*

(Note: hypotheses (iii) and (iv) ensure that $\rho$ is potentially semistable.)

Since the appearance of [D1] other refinements and generalizations of this theorem have appeared, most notably in [CDT], [BCDT], and [Di], but to simplify both notation and statements we state only the theorem above.

Subsequent efforts have also gone a long way towards allowing $\mathbf{Q}$ to be replaced by a totally real field $F$ in Theorem 1 and towards loosening condition (v). In order to explain these efforts and the obstacles encountered, we first recall the strategy behind the proof of Theorem 1.

## 3. The strategy (very briefly)

Suppose $p$ is an odd prime and suppose $\rho$ satisfies the hypotheses of Theorem 1. Replace $\rho$ by a conjugate taking values in some $\mathrm{GL}_2(\mathcal{O})$ as was done to define $\overline{\rho}$. Let $k$ be the residue field of $\mathcal{O}$.

*Step 1. Formulate deformation problems*

By an $\mathcal{O}$-deformation of $\overline{\rho}$ we mean an equivalence class of continuous representations $\sigma : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(A)$, $A$ a complete local Noetherian $\mathcal{O}$-algebra with residue field $k$ and maximal ideal $\mathfrak{m}_A$, such that $\rho \bmod \mathfrak{m}_A = \overline{\rho}$.

Let $\Sigma_0$ be the set of primes at which $\overline{\rho}$ is ramified. Given a finite set of primes $\Sigma \supseteq \Sigma_0$, we say that an $\mathcal{O}$-deformation $\sigma$ is *of type-*$\Sigma$ provided

- $\sigma$ is unramified away from $\Sigma \cup \{p\}$,
- if $\rho$ is nearly ordinary at $p$ then $\sigma|_{D_p} \cong \begin{pmatrix} \psi_1 & * \\ & \psi_2 \end{pmatrix}$ with $\overline{\psi}_2 = \overline{\chi}_2$,
- if $\rho$ is flat at $p$, but not nearly ordinary, then for all $n$, $\sigma \bmod \mathfrak{m}_A^n$ arises from a finite flat group scheme over $\mathbf{Z}_p$ with an $A$-action.

Clearly $\rho$ is of type-$\Sigma$ for any $\Sigma$ containing all the primes at which $\rho$ is ramified.

We call an $\mathcal{O}$-deformation $\sigma$ *minimal* if

- $\sigma$ is of type-$\Sigma_0$ with the slightly stronger condition that
- if $\rho$ is flat at $p$, then for all $n$, $\sigma \bmod \mathfrak{m}_A^n$ arises from a finite flat group scheme over $\mathbf{Z}_p$ with an $A$-action.
- $\sigma$ is "minimally ramified at each $\ell \in \Sigma_0$."

The third condition (which we do not make precise) is such that the ramification in $\sigma|_{D_\ell}$ is essentially the same as that of $\overline{\rho}|_{D_\ell}$.

From the theory of deformations of Galois representations as introduced by Mazur [M1], [M2] it follows that there are *universal* $\mathcal{O}$-deformations

$$\rho_\Sigma : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(R_\Sigma) \quad \text{and} \quad \rho_{\min} : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(R_{\min})$$

of type-$\Sigma$ and minimal, respectively. The universality is such that if $\sigma$ is any $\mathcal{O}$-deformation of type-$\Sigma$ into $\mathrm{GL}_2(A)$ then there exists a unique map $R_\Sigma \to A$ of local $\mathcal{O}$-algebras such that $\sigma$ is obtained by combining this map with $\rho_\Sigma$, and similarly for minimal deformations.

*Step 2. Find some modular deformations*

Let $\Pi_\Sigma$ and $\Pi_{\min}$ be the sets of cuspidal automorphic representations $\pi$ such that $\rho_{\pi,p}$ is isomorphic to an $\mathcal{O}$-deformation that is of type-$\Sigma$ or minimal (with scalars extended to $\overline{\mathbf{Q}}_p$, of course), respectively. Let $\mathbf{T}_\Sigma$ be the $\mathcal{O}$-subalgebra of $\prod_{\pi \in \Pi_\Sigma} \overline{\mathbf{Q}}_p$ generated by $T_\ell = (\mathrm{trace}\, \rho_{\pi,p}(\mathrm{frob}\, \ell))_{\pi \in \Pi_\Sigma}$, $\ell \notin \Sigma \cup \{p\}$. Let $\mathbf{T}_{\min}$ be defined similary but with $\Pi_\Sigma$ replaced by $\Pi_{\min}$. Then both $\mathbf{T}_\Sigma$ and $\mathbf{T}_{\min}$ are complete, local Noetherian $\mathcal{O}$-algebras with residue field $k$. When $\rho$ is nearly ordinary at $p$ but not flat, then $\mathbf{T}_\Sigma$ is a nearly-ordinary Hecke ring in the sense of Hida [H1] and hence 3-dimensional. Otherwise these rings are finite $\mathcal{O}$-algebras. That the set $\Pi_{\min}$ is non-empty (and so $\mathbf{T}_{\min}$ and $\mathbf{T}_\Sigma$ are non-zero) is a consequence of hypothesis (vi) of Theorem 1 and the proof of Serre's epsilon conjecture (see [D2]).

There are $\mathcal{O}$-deformations

$$\rho_\Sigma^{\mathrm{mod}} : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{T}_\Sigma) \quad \text{and} \quad \rho_{\min}^{\mathrm{mod}} : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{T}_{\min})$$

of type-$\Sigma$ and minimal, respectively, such that

$$\text{(2.1a)} \qquad \operatorname{trace} \rho_\Sigma^{\mathrm{mod}}(\operatorname{frob}_\ell) = T_\ell, \quad \ell \notin \Sigma \cup \{p\},$$

and

$$\text{(2.1b)} \qquad \operatorname{trace} \rho_{\min}^{\mathrm{mod}}(\operatorname{frob}_\ell) = T_\ell, \quad \ell \notin \Sigma_0 \cup \{p\}.$$

Thus there are surjective maps

$$\phi_\Sigma : R_\Sigma \to \mathbf{T}_\Sigma \quad \text{and} \quad \phi_{\min} : R_{\min} \to \mathbf{T}_{\min}$$

coming from the universality of $R_\Sigma$ and $R_{\min}$ (with surjectivity coming from (2.1a,b)).

*Step 3. Prove that $\phi_{\min}$ is an isomorphism*

This is essentially done by realizing both $R_{\min}$ and $\mathbf{T}_{\min}$ as quotients of the same power series ring by the same ideal, at the same time showing that they are also both complete intersection rings. The power series rings and the quotient maps are obtained by "patching" together various quotients of auxiliary deformation rings.

*Step 4. Deduce from Step 2 that $\phi_\Sigma$ is an isomorphism*

This is done by comparing two numerical invariants. By extending $\mathcal{O}$ if necessary, we can assume that there exists a $\pi$ such that $\rho_{\pi,p}$ comes from a minimal deformation into $\mathrm{GL}_2(\mathcal{O})$ (i.e, is isomorphic to such a representation with scalars extended to $\overline{\mathbf{Q}}_p$). This gives rise to maps

$$\psi_{\min} : \mathbf{T}_{\min} \to \mathcal{O} \quad \text{and} \quad \psi_\Sigma : \mathbf{T}_\Sigma \to \mathcal{O}$$

as well as maps

$$\theta_{\min} : R_{\min} \to \mathcal{O} \quad \text{and} \quad \theta_\Sigma : R_\Sigma \to \mathcal{O}.$$

Note that $\theta_{\min} = \psi_{\min} \circ \phi_{\min}$, $\theta_\Sigma = \psi_\Sigma \circ \phi_\Sigma$, and $\psi_\Sigma$ and $\theta_\Sigma$ factor through the natural surjective maps $\mathbf{T}_\Sigma \to \mathbf{T}_{\min}$ and $R_\Sigma \to R_{\min}$, respectively. Put

$$\eta_{\min} = \phi_{\min}(\mathrm{Ann}_{\mathbf{T}_{\min}}(\ker \phi_{\min})) \quad \text{and} \quad \eta_\Sigma = \phi_\Sigma(\mathrm{Ann}_{\mathbf{T}_\Sigma}(\ker \phi_\Sigma)).$$

The isomorphism of Step 3, being an isomorphism of complete intersection rings, implies (cf. [DRS]) that

$$\text{(2.2)} \qquad \# \ker(\theta_{\min})/(\ker(\theta_{\min}))^2 = \#\mathcal{O}/\eta_{\min}.$$

That $\phi_\Sigma$ is an isomorphism would follow from

$$\text{(2.3)} \qquad \# \ker(\theta_\Sigma)/(\ker(\theta_\Sigma))^2 = \#\mathcal{O}/\eta_\Sigma.$$

To establish (2.3) one needs to give an upper bound on the difference of the left hand sides of (2.2) and (2.3) and a lower bound on the difference of the right hand sides and hope that these bounds are the same. The left hand sides are orders of global Galois cohomology groups, but their difference is bounded from above by a product of orders of some local Galois cohomology groups that can be easily computed using local Tate duality. The right hand sides essentially measure congruences between cuspforms and the difference between the two sides can be bounded below by generalizations of Ribet's methods (see [W1, Chapter 2]). This last step requires some technical results on the injectivity of various cohomology groups of congruence subgroups and their freeness over the Hecke rings. This latter result is a consequence of the "multiplicity one" result in Chapter 2 of [W1] (nowadays it can be proven in other ways - but that's the next section).

*Step 5. Deduce that $\rho$ is modular*

Suppose $\Sigma$ contains all the primes at which $\rho$ is ramified. The map $\mathbf{T}_\Sigma \cong R_\Sigma \to \mathcal{O}$ corresponding to $\rho$ via the universality of $R_\Sigma$ sends $T_\ell$ to $\operatorname{trace}\rho(\operatorname{frob}_\ell)$ for all $\ell \notin \Sigma$ by (2.1b). From this one easily deduces that $\rho \cong \rho_{\pi,p}$ for some $\pi \in \Pi_\Sigma$ (in the nearly ordinary case one needs the main results of [H1]).

## 4. Residually irreducible representations over totally real fields

About a year after the publication of [W1], K. Fujiwara released a preprint [F1] that generalized much of Steps 1-3 of the preceding strategy to totally real fields.

**Theorem 2** ([F1]). *Suppose $p$ is an odd prime and $F$ is a totally real field. If $\rho : \operatorname{Gal}(\overline{F}/F) \to \operatorname{GL}_2(\overline{\mathbf{Q}}_p)$ is a continuous representation such that*

- (i) *$\rho$ is irreducible and unramified at all but finitely many places,*
- (ii) *$\rho$ is odd,*
- (iii) *$\det\rho = \chi\varepsilon^m$ with $\chi$ finite and $m \geq 1$,*
- (iv) *either $p$ is totally unramified in $F$ and $\rho$ is flat at all $v|p$ or $\rho$ is not flat at any $v|p$ but is nearly ordinary and $D_v$-distinguished at all $v|p$,*
- (v) *$\overline{\rho}|_{\operatorname{Gal}(\overline{F}/F(\zeta_p))}$ is irreducible,*
- (vi) *$\overline{\rho} \cong \overline{\rho}_{\pi,p}$ for some $\pi$ such that $\rho_{\pi,p}$ comes from a minimal deformation of $\overline{\rho}$,*
- (vii) *$\rho$ comes from a minimal deformation of $\overline{\rho}$,*

*then $\rho \in \mathcal{MG}_{F,p}$.*

Actually, Fujiwara proves a stronger result. In particular, he proves a result analogous to Step 3 above: he identifies a deformation ring with a Hecke ring.

Step 1 of the strategy sketched in §2 generalizes to the setting of Theorem 2 without any difficulty. Step 2 also generalizes, but we also need to know that $\Pi_{\min}$ is not empty. This is ensured by hypothesis (vi) of Theorem 2 which is much stronger than the corresponding hypothesis of Theorem 1.

A stronger hypothesis is needed since we do not have a version of Serre's $\varepsilon$-conjecture for totally real fields (but see [J], [Ra], and [F3] for progress towards such).

One of the most significant new ideas in [F1] comes in generalizing Step 3. The argument in [TW] relies on the fact that certain $\mathbf{T}_{\min}$-modules are free. This fact is proven in [W1] by a delicate geometric argument that does not seem to generalize to a general totally real field. But Fujiwara discovered that by "patching" both modules and rings he could simultaneously prove that $R_{\min}\cong\mathbf{T}_{\min}$ *and* that the $\mathbf{T}_{\min}$-modules are free (a fact which is important for Step 4). This was independently discovered by F. Diamond [D3].

Step 4 has yet to be completely generalized, whence hypothesis (vii) of Theorem 2. However in [F2], a refinement of [F1], Fujiwara has made great progress towards doing so.

Some of the problems encountered in attempting to generalize the strategy can be avoided if one is content to merely prove $\rho \in \mathcal{MG}_{F,p}$ and not $R_\Sigma\cong\mathbf{T}_\Sigma$. This is thanks to Langlands' results on the image of solvable base change which imply that $\rho \in \mathcal{MG}_{F,p}$ if $\rho \in \mathcal{MG}_{L,p}$ for some finite, totally real extension $L$ of $F$ having solvable normal closure over $F$. This can be exploited as follows. The main result of [SW3] essentially asserts that $\overline{\rho}$ being irreducible combined with $\overline{\rho}\cong\overline{\rho}_{\pi,p}$ for some $\pi$ such that $\rho_{\pi,p}$ comes from a deformation that is minimal at $p$ (meaning that it satisfies the conditions of a minimal deformation for all $v|p$ but not necessarily at other places) implies hypothesis (vi) of Theorem 2 over some extension $L$ of $F$ as above. Combining this with Fujiwara's results and the observations about base-change should yield the following theorem.

**"Theorem 3"**. *Suppose $p$ is an odd prime and $F$ is a totally real field. If $\rho : \mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ is a continuous representation such that*

(i) *$\rho$ is irreducible and unramified at all but finitely many places,*
(ii) *$\rho$ is odd,*
(iii) *$\det\rho = \chi\varepsilon^m$ with $\chi$ finite and $m \geq 1$,*
(iv) *either $p$ is totally unramified in $F$ and $\rho$ is flat at all $v|p$ or $\rho$ is not flat at any $v|p$ but is nearly ordinary and $D_v$-distinguished at all $v|p$,*
(v) *$\overline{\rho}|_{\mathrm{Gal}(\overline{F}/F(\zeta_p))}$ is irreducible,*
(vi) *$\overline{\rho}\cong\overline{\rho}_{\pi,p}$ for some $\pi$ such that $\rho_{\pi,p}$ comes from a deformation of $\overline{\rho}$ that is minimal at $p$,*

*then $\rho \in \mathcal{MG}_{F,p}$.*

We use "should" and put quotes around the label because a proof has not yet been written down.

Exploiting a similar observation Wiles and the author were able to loosen hypothesis (v) of Theorem 1 in some cases.

**Theorem 4** ([SW4]). *Suppose $p$ is an odd prime and $F$ is a totally real field. If $\rho : \mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ is a continuous representation such that*

(i) *$\rho$ is irreducible and unramified at all but finitely many places,*
(ii) *$\rho$ is odd,*
(iii) $\det\rho = \chi\varepsilon^m$ *with $\chi$ finite and $m \geq 1$,*
(iv) *$\rho$ is nearly ordinary at $v$ and $D_v$-distinguished for all $v|p$,*
(v) *$\overline{\rho}$ is irreducible,*
(vi) *$\overline{\rho} \cong \overline{\rho}_{\pi,p}$ for some $\pi$ such that $\rho_{\pi,p}$ is nearly ordinary at each $v|p$ and $\chi_2$-good,*

*then $\rho \in \mathcal{MG}_{F,p}$.*

The condition of being $\chi_2$-good is that if $\rho|_{D_v} \cong \begin{pmatrix} \chi_{1,v} & * \\ & \chi_{2,v} \end{pmatrix}$ then

$$\rho_{\pi,p}|_{D_v} \cong \begin{pmatrix} \psi_{1,v} & * \\ & \psi_{2,v} \end{pmatrix}$$

with $\overline{\psi}_{2,v} = \overline{\chi}_{2,v}$ for all $v|p$. (This is an unnecessary hypothesis if $F = \mathbf{Q}$ or if no $\overline{\rho}|_{D_v}$ is split.)

The proof of Theorem 4 is an adaptation of the methods employed in [SW2] to handle cases where $\overline{\rho}$ is reducible, so we will reserve comments for that case, which is discussed in the next section.

## 5. Residually reducible representations

So far all the results discussed have included the condition that $\overline{\rho}$ be irreducible. What about the case where $\overline{\rho}$ is reducible? Such representations were the focus of [SW1] and [SW2]. In [SW1] some special cases where one could identify the deformation rings with Hecke rings were examined; roughly the cases where $\overline{\rho} = 1 \oplus \chi$ and there is a unique extension of 1 by $\chi$. A different tack was taken in [SW2] that yielded the following theorem.

**Theorem 5** ([SW2]). *Suppose $p$ is an odd prime and $F$ is a totally real field. Suppose also that $\rho : \mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ is a continuous representation. If*

(i) *$\rho$ is irreducible and unramified at all but finitely many places,*
(ii) *$\rho$ is odd,*
(iii) $\det\rho = \chi\varepsilon^m$ *with $\chi$ finite and $m \geq 1$,*
(iv) *$\overline{\rho} = \chi_1 \oplus \chi_2$ and $(\chi_1/\chi_2)|_{D_v} \neq 1$ for all $v|p$,*
(v) *for all $v|p$, $rho|_{D_v} = \begin{pmatrix} \psi_{1,v} & * \\ & \psi_{2,v} \end{pmatrix}$, $\bar{\psi}_{2,v}|_{D_v} = \chi_2$, and $\psi_{2,v}|_{I_v}$ has finite order,*
(vi) *$F(\chi_1/\chi_2)$ is abelian over $\mathbf{Q}$,*

*then $\rho \in \mathcal{MG}_{F,p}$.*

This actually follows from a theorem for general $F$ that requires the existence of solvable extension of $F(\chi_1/\chi_2)$ (which is the splitting field over $F$ of the character $\chi_1/\chi_2$) whose class group has $p$-rank relatively small in comparison to its degree. When $F(\chi_1/\chi_2)$ is abelian this is known thanks to a theorem of Washington [Wa].

We now sketch the main ingredients that go into the proof of Theorem 5.

We assume $\rho$ satisfies the hypotheses of Theorem 5. Replacing $\rho$ by its twist by the Teichmüller lift of $\chi_2^{-1}$, we can assume

- $\rho : \mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}_2(\mathcal{O})$
- $\overline{\rho} = 1 \oplus \chi \in \mathrm{GL}_2(k)$, $\chi|_{D_v} \neq 1$ for all $v|p$.

1. *Deformation rings*

For any finite set $\Sigma$ of finite places of $F$ containing all those dividing $p$ and all those at which $\chi$ is ramified, let $G_\Sigma$ be the Galois group of the maximal extension of $F$ unramified outside of $\Sigma$ and the places above infinity. Let

$$H_\Sigma = \ker\{H^1(G_\Sigma, k(\chi^{-1})) \to \prod_{v|p} H^1(D_v, k(\chi^{-1}))\},$$

where the maps are the restriction maps. For each $0 \neq c \in H_\Sigma$ there is a non-split representation $\rho_c : \mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}_2(k)$ such that

- $\rho_c$ is unramified at all finite places not in $\Sigma$,
- $\rho_c = \left(\begin{smallmatrix} 1 & * \\ & \chi \end{smallmatrix}\right)$ with $c$ being the cohomology class associated to $*$.
- $\rho_c|_{D_v} = 1 \oplus \chi \in \mathrm{GL}_2(k)$ for all $v|p$.

We will say that an $\mathcal{O}$-deformation $\sigma$ of $\rho_c$ is *of type*-$(c, \Sigma)$ if

- $\sigma$ is unramified at all finite places not in $\Sigma$,
- $\sigma|_{D_v} \cong \begin{pmatrix} \psi_{1,v} & * \\ & \psi_{2,v} \end{pmatrix}$ with $\overline{\psi}_{1,v} = \chi$ for all $v|p$.

There is a universal deformation $\rho_{c,\Sigma} : \mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}_2(R_{c,\Sigma})$ of type-$(c, \Sigma)$.

2. *Hecke rings*

Define $\mathbf{T}_\Sigma$ and $T_v$ ($v$ a finite place not in $\Sigma$) as in Step 2 of §2, but now let $\Pi_\Sigma$ be the set of $\pi$'s such that $\rho_{\pi,p}$ comes from a deformation of type-$(c, \Sigma)$ for *some* $0 \neq c \in H_\Sigma$. Unfortunately, in general there is no longer a natural representation into $\mathrm{GL}_2(\mathbf{T}_\Sigma)$. Essentially this is because $\dim_k H_\Sigma$ can be bigger than one. However, one does have a pseudo-representation into $\mathbf{T}_\Sigma$.

3. *Pseudo-deformations*

By an ($\mathcal{O}$-) *pseudo-deformation of type*-$\Sigma$ we mean a complete local Noetherian $\mathcal{O}$-algebra $A$ with residue field $k$ and a triple $\phi = (a, d, x)$ of continuous maps $a, d : G_\Sigma \to A$ and $x : G_\Sigma \times G_\Sigma \to A$ such that

- $a \bmod \mathfrak{m}_A = 1$, $d \bmod \mathfrak{m}_A = \chi$, $x \bmod \mathfrak{m}_A = 0$,
- $a(g_1g_2) = a(g_1)a(g_2) + x(g_1, g_2)$,
- $d(g_1g_2) = d(g_1)d(g_2) + x(g_2, g_1)$,
- $x(g_1, g_2)x(g_3, g_4) = x(g_1, g_4)x(g_3, g_2)$,
- $\cdots$.

and some other equally illuminating relations (see Section 2.4 of [SW2]). Essentially these capture the traces of $\mathcal{O}$-deformations of any $\rho_c$. If

$$\sigma(g) = \left(\begin{smallmatrix} a_g & b_g \\ c_g & d_g \end{smallmatrix}\right) \in \mathrm{GL}_2(A)$$

is an $\mathcal{O}$-deformation of some $\rho_c$ (suitably normalized with respect to some fixed complex conjugation), then

$$a(g) = a_g, \quad d(g) = d_g, \quad x(g_1, g_2) = b(g_1)c(g_2) \tag{4.1}$$

defines a pseudo-deformation of type-$\Sigma$. In particular, the universal deformation $\rho_{c,\Sigma}$ determines a pseudo-deformation $\phi_{c,\Sigma}$ of type-$\Sigma$.

It turns out that there is a universal pseudo-deformation of type-$\Sigma$, call it $(R_\Sigma^{ps}, \phi_\Sigma^{ps})$. Also, there is a pseudo-deformation $\phi_\Sigma^{\mathrm{mod}} = (a^{\mathrm{mod}}, d^{\mathrm{mod}}, x^{\mathrm{mod}})$ into $\mathbf{T}_\Sigma$ such that

$$a^{\mathrm{mod}}(\mathrm{frob}_v) + d^{\mathrm{mod}}(\mathrm{frob}_v) = T_v. \tag{4.2}$$

Thus there are maps

$$r_{c,\Sigma} : R_\Sigma^{ps} \to R_{c,\Sigma} \quad \text{and} \quad r_\Sigma : R_\Sigma^{ps} \to \mathbf{T}_\Sigma$$

coming from the universality of $R_\Sigma^{ps}$.

4. *Pro-modular primes*

We define a prime $\mathfrak{p}$ of $R_{c,\Sigma}$ to be *pro-modular* if there exists a map $\phi : \mathbf{T}_\Sigma \to R_{c,\Sigma}/\mathfrak{p}$ such that

$$\phi \circ r_\Sigma = \pi_{\mathfrak{p}} \circ r_{c,\Sigma}$$

where $\pi_{\mathfrak{p}}$ is the natural surjection $R_{c,\Sigma} \to R_{c,\Sigma}/\mathfrak{p}$. In [SW2] we show that under suitable hypotheses every prime of each $R_{c,\Sigma}$ is promodular. This is done by a delicate induction argument including a generalization of the strategy sketched in §2.

5. *Finishing up*

To show that $\rho \in \mathcal{MG}_{F,p}$, we first note that we may assume that $\rho$ is a deformation of type-$(c, \Sigma)$ for some $c$. Then it follows from Step 4 that the kernel of the corresponding map $R_{c,\Sigma} \to \mathcal{O}$ is pro-modular. From this and (4.1) and (4.2) one can deduce that there is a map $\mathbf{T}_\Sigma \to \mathcal{O}$ such that $T_v$ maps to trace $\rho(\mathrm{frob}_v)$ for each $v \notin \Sigma$. And from this it follows that $\rho \cong \rho_{\pi,p}$ for some $\pi$ just as in Step 5 of §2.

## 6. Odds and Ends

In addition to the results high-lighted above, there have been a number of other refinements, generalizations, and additions to the circle of ideas that started with [W1]. We briefly discuss a few.

$p = 2$. M. Dickinson [Di] has begun to remove the condition that $p$ be odd.

*Artin's Conjecture.* This Artin's Conjecture is that the Artin $L$-function attached to an irreducible Galois representations $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{C})$ has an analytic continuation to the whole complex plane. The ground-breaking work of Langlands and Tunnell in the case where the image of this representation is solvable played a fundamental role in [W1] and subsequent work. This in turn has led to some partial results for the last remaining case (when the projective image is icosahedral). See [BDST], [BT] for example.

$\mathbf{Q}$-*curves.* The most spectacular application of the results stemming from [W1] has been the proof that all elliptic curves over $\mathbf{Q}$ are modular. While not all elliptic curves (over a number field) are expected to be modular (in the sense of being a quotient of the jacobian of a modular curve), there is a class that is: the $\mathbf{Q}$-curves. An elliptic curve over a Galois extension is called a $\mathbf{Q}$-*curve* if it is isogeneous to all of it's Galois conjugates. In [ES] results mentioned in previous sections are used to prove the modularity of many such curves.

$n > 2$. There has also been effort to generalize the aforementioned theorems by replacing $\mathrm{GL}_2$ with $\mathrm{GL}_n$ and $\pi$ with an automorphic representation on $\mathrm{GL}_{n_{/F}}$. Some partial success has been achieved by M. Harris and R. Taylor. Generalizing all the steps in §2 poses series technical difficulties, not the least of which is finding the Galois representations associated to the $\pi$'s on $\mathrm{GL}_{n_{/F}}$!

*Alternate approaches.* C. Khare and R. Ramakrishna (unpublished) have developed an alternate approach to identifying deformation rings with Hecke rings (and thereby establishing modularity) that yields many cases of Theorem 1. Their approach is fairly different, and among other things avoids the "patching arguments" usually employed in Step 3 of §2. To do justice to their novel ideas would really require another lecture.

## References

[BR] D. Blasius, J. Rogawski, *Motives for Hilbert modular forms.* Invent. Math. **114** (1993), no. 1, 55–87.

[BCDT] C. Breuil, B. Conrad, F. Diamond, R. Taylor, *On the modularity of elliptic curves over* $\mathbf{Q}$*: wild 3-adic exercises.* J. Amer. Math. Soc. **14** (2001), no. 4, 843–939.

[BDST] K. Buzzard, M. Dickinson, N. Shepherd-Barron, R. Taylor, *On icosahedral Artin representations.* Duke Math. J. **109** (2001), no. 2, 283–318.

[BT] K. BUZZARD, R. TAYLOR, *Companion forms and weight one forms.* Ann. of Math. **149** (1999), no. 3, 905–919.

[CDT] B. CONRAD, F. DIAMOND, R. TAYLOR, *Modularity of certain potentially Barsotti-Tate Galois representations.* J. Amer. Math. Soc. **12** (1999), no. 2, 521–567.

[DRS] B. DE SMIT, K. RUBIN, R. SCHOOF, *Criteria for complete intersections.* In: Modular forms and Fermat's Last Theorem, 343–356, Springer, 1997.

[D1] F. DIAMOND, *On deformation rings and Hecke rings.* Ann. of Math. (2) **144** (1996), no. 1, 137–166.

[D2] F. DIAMOND, *The refined conjecture of Serre.* In: Elliptic Curves, Modular forms, and Fermat's Last Theorem (ed. J. Coates), International Press, Cambridge, MA, 1995.

[D3] F. DIAMOND, *The Taylor-Wiles construction and multiplicity one.* Invent. Math. **128** (1997), no. 2, 379–391.

[Di] M. DICKINSON, *On the modularity of certain 2-adic Galois representations.* Duke Math. J. **109** (2001), no. 2, 319–382.

[ES] J. ELLENBORG, C. SKINNER, *On the modularity of* $\mathbf{Q}$*-curves.* Duke Math. J. **109** (2001), no. 1, 97–122.

[Fo] J.-M. FONTAINE, *Représentations $\ell$-adiques potentiellement semi-stables.* In: Périodes $p$-adiques, Asterisque **223** (1994) 321–247.

[FoM] J.-M. FONTAINE, B. MAZUR, *Geometric Galois representations.* In: Elliptic Curves, modular forms, and Fermat's Last Theorem (Hong Kong, 1993), pp. 41–78, Internat. Press, 1995.

[F1] K. FUJIWARA, *Deformation rings and Hecke algebras in the totally real case*, preprint (1996).

[F2] K. FUJIWARA, *Deformation rings and Hecke algebras in the totally real case*, preprint (1999).

[F3] K. FUJIWARA, it Level optimazation in the totally real case, preprint (1999).

[H1] H. HIDA, *On nearly ordinary Hecke algebras for $GL(2)$ over totally real fields.* In: Algebraic number theory, Adv. Stud. Pure Math. **17**, 139–169, Academic Press, 1989.

[H2] H. HIDA, *Nearly ordinary Hecke algebras and Galois representations of several variables.* In: Algebraic analysis, geometry, and number theory (Baltimore, MD 1988), 115–138, John Hopkins Univ. Press, 1989.

[J] F. JARVIS, *Level lowering for modular mod $\ell$ representations over totally real fields.* Math. Ann. **313** (1999), 141–160.

[M1] B. MAZUR, *Deforming Galois representations.* In: Galois Groups over $\mathbf{Q}$, vol. 16, MSRI Publications, Springer, 1989.

[M2] B. MAZUR, *An introduction to the deformation theory of Galois representations.* In: Modular Forms and Fermat's Last Theorem (eds. G. Cornell et al.), Springer-Verlag, New York, 1997.

[Ra] A. RAJAEI, *On lowering the levels in modular mod $\ell$ Galois representations of totally real fields.* Thesis, Princeton University, 1998.

[SW1] C. SKINNER, A. WILES, *Ordinary representations and modular forms.* Proc. Nat. Acad. Sci. U.S.A. **94** (1997), no. 20, 10520–10527.

[SW2] C. SKINNER, A. WILES, *Modular forms and residually reducible representations.* Publ. Math. IHES **89** (1999), 5–126.

[SW3] C. SKINNER, A. WILES, *Base change and a problem of Serre.* Duke Math. J. **107** (2001), no. 1, 15–25.

[SW4] C. SKINNER, A. WILES, *Nearly ordinary deformations of irreducible residual representations.* Ann. Fac. Sci. Toulouse Math. (6) **10** (2001), no. 1, 185–215.

[Ta] J. TATE, *Number theoretic background.* In: Automorphic forms, representations and $L$-functions, Part 2, Proc. Sympos. Pure Math., XXXIII, pp. 3–26, Amer. Math. Soc., Providence, R.I., 1979.

[T] R. TAYLOR, *On Galois representations associated to Hilbert modular forms.* In: Elliptic Curves, Modular forms, and Fermat's Last Theorem (ed. J. Coates), International Press, Cambridge, MA, 1995.

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

[Wa] L. WASHINGTON, *The non-p-part of the class number in a cyclotomic* $\mathbf{Z}_p$*-extension*. Invent. Math. **49** (1978), no. 1, 87–97.

[W1] A. WILES, *Modular elliptic curves and Fermat's Last Theorem*. Ann. of Math. (2) **142** (1995), 443–551.

[W2] A. WILES *On ordinary* $\lambda$*-adic representations associated to modular forms*. Invent. Math. **94** (1988), no. 3, 529–573.



Chris SKINNER
Department of Mathematics
University of Michigan
Ann Arbor, MI 48109
USA
*E-mail* : cskinner@umich.edu