AN EXTENSION OF WILES' RESULTS

FRED DIAMOND

1. Introduction

Suppose that E is an elliptic curve defined over $\mathbf{Q}$. We wish to prove that E is modular, or equivalently, that the associated ℓ-adic representation

$$\rho : G_{\mathbf{Q}} \to GL_2(\mathbf{Z}_\ell)$$

is modular for some prime ℓ.

If we are assuming that E is semistable, i.e., has square-free conductor, then we can impose some convenient hypotheses on the local behavior of the Galois representations ρ we consider. By "local behavior," we mean the behavior of the representation

$$\theta : G_p \to GL_2(\mathbf{Z}_\ell)$$

defined by restricting ρ to a decomposition group at p.

Recall that if E has good reduction at p, then θ is unramified. If E has multiplicative reduction at p, then a convenient description of θ results from the Tate parametrization of E (§17 of [S]). In particular, we see that

$$\theta|_{I_p} \sim \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}.$$

To consider elliptic curves with additive reduction at some primes $p \neq \ell$, we must allow more general types of θ. We can actually consider representations ρ with arbitrary local behavior at primes $p \neq \ell$. This is carried out in [D] where, building on the work of Wiles [W] and Taylor-Wiles [TW], we prove a result of the form

$$\bar{\rho} \text{ modular } \Rightarrow \ \rho \text{ modular }, \tag{1}$$

and deduce

Theorem 1.1. *If E has good or multiplicative reduction at 3 and 5, then E is modular.*

The details of the proof can be found in [D]. Here we give an exposition which we hope is more motivated and systematic. We often follow [DDT], admitting results which are straightforward generalizations of those there or elsewhere in this volume. For the proofs of some of the key lemmas, we refer completely to [D].

This article is structured as follows:

Some background is given in §2 on local Galois representations

$$\sigma : G_p \to GL_2(k),$$

where $p \neq \ell$ and k is an algebraic closure of $\mathbf{F}_\ell$. A classification of the possible σ, though not logically necessary for the proof of theorem 1.1, helps provide some insight into local Galois representations and their deformations. (The appendix with K. Kramer determines precisely how local Galois representations arising from elliptic curves fit into this classification.)

In §3 we explain what it means for a deformation of σ to be "minimally ramified" at p.

Suppose that ℓ is odd and

$$\bar{\rho} : G_{\mathbf{Q}} \to GL_2(k)$$

is an irreducible representation which is semistable at ℓ. We formulate in §4 a certain deformation problem for each finite set of primes Σ. This deformation problem turns out to be representable by a ring R_Σ whose tangent space is described in terms of Galois cohomology (see [M]).

Suppose now that $\bar{\rho}$ is modular. The goal of §5 is to define a corresponding Hecke algebra $\mathbf{T}_\Sigma$, and modular deformation

$$\tau : G_{\mathbf{Q}} \to GL_2(\mathbf{T}_\Sigma)$$

arising from a homomomorphism $\phi_\Sigma : R_\Sigma \to \mathbf{T}_\Sigma$ (see [Ri2]).

If we can show that it is an isomorphism, then we obtain a result of the form (1) as a corollary. The main results are stated in §6.

To prove that ϕ_Σ is an isomorphism, we must modify some of the techniques used in [W] and [TW]. In particular, the analysis of the Hecke algebras becomes more difficult. In our sketch of the proof in §7, we indicate where the complications arise, but give only a rough idea of how they are dealt with in [D].

2. Local representations mod ℓ

Suppose that p is a prime. We let $G_p = \mathrm{Gal}\,(\bar{\mathbf{Q}}_p/\mathbf{Q}_p)$ and let I_p denote the inertia subgroup of G_p. Suppose that ℓ is an odd prime different from p and consider continuous representations

$$\sigma : G_p \to \mathrm{GL}_2(k),$$

where k is an algebraic closure of $\mathbf{F}_\ell$. We let $\tilde{\sigma}$ denote the associated projective representation.

We let χ denote the cyclotomic character $G_p \to k^\times$. Note that χ is nontrivial if and only if $p \not\equiv 1 \bmod \ell$. In that case, we write sp_2 for the

representation

$$\begin{pmatrix} \chi & u \\ 0 & 1 \end{pmatrix}, \tag{2}$$

where u is a cocycle representing the image of a uniformizer under the Kummer map

$$\mathbf{Q}_p^\times \to \mathbf{Q}_p^\times/(\mathbf{Q}_p^\times)^\ell \to H^1(G_p, k(1)).$$

The equivalence class is independent of the choice of uniformizer and cocycle. (See the proof of proposition 2.2 below.)

If ψ is a character $G_p \to k^\times$, then we write $k(\psi)$ for the one-dimensional vector space over k on which G_p acts via ψ. Recall that two representations σ_1 and σ_2 are called twist-equivalent if σ_1 is equivalent to $\psi \otimes \sigma_2$ for some character $\psi : G_p \to k^\times$.

We classify σ according to the following four types of behavior (principal, special, vexing or harmless).

P : σ is reducible and $\sigma|_{I_p}$ is decomposable.
S : σ is reducible and $\sigma|_{I_p}$ is indecomposable.
V : σ is irreducible and $\sigma|_{I_p}$ is reducible.
H : σ is irreducible and $\sigma|_{I_p}$ is irreducible.

Proposition 2.1. *The following are equivalent:*

1. *σ is reducible and $\sigma|_{I_p}$ is decomposable.*
2. *σ is twist-equivalent to a representation either of the form*
   (a) $\begin{pmatrix} \psi & 0 \\ 0 & 1 \end{pmatrix}$ *for some character ψ, or*
   (b) $\begin{pmatrix} 1 & \phi \\ 0 & 1 \end{pmatrix}$ *for some additive unramified character ϕ.*
3. *Either $\tilde{\sigma}(G_p)$ is cyclic of order not divisible by ℓ, or it has order ℓ and $\tilde{\sigma}(I_p)$ is trivial.*
4. *$\tilde{\sigma}(G_p)$ is cyclic and the order of $\tilde{\sigma}(I_p)$ is not divisible by ℓ.*

Proof: Suppose 1 holds. Then σ is twist-equivalent to a representation of the form

$$\begin{pmatrix} \psi & u \\ 0 & 1 \end{pmatrix}$$

for some character ψ, where u is a cocycle representing a class

$$x \in H^1(G_p, k(\psi)).$$

If σ is indecomposable, then x is nontrivial. On the other hand, the image of x in $H^1(I_p, k(\psi))$ vanishes, so x is in the image of $H^1(G_p/I_p, k(\psi)^{I_p})$. This last group is trivial unless ψ is trivial, so 2 follows.

The implications $2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 1$ are clear.

Proposition 2.2. *The following are equivalent:*

1. *σ is reducible and $\sigma|_{I_p}$ is indecomposable.*
2. *$p \equiv 1 \bmod \ell$ and σ is twist-equivalent to a representation of the form*

$$\begin{pmatrix} 1 & \phi \\ 0 & 1 \end{pmatrix}$$

for some additive ramified character ϕ, or $p \not\equiv 1 \bmod \ell$ and σ is twist-equivalent to sp_2.
3. *$p \equiv 1 \bmod \ell$, $\tilde{\sigma}(I_p)$ has order ℓ and $\tilde{\sigma}(G_p)$ has order dividing ℓ^2, or $p \not\equiv 1 \bmod \ell$, $\tilde{\sigma}(I_p)$ has order ℓ and $\tilde{\sigma}(G_p)$ has order $d\ell$ where d is the order of p in $\mathbf{F}_\ell^\times$.*
4. *$\tilde{\sigma}(I_p)$ is cyclic of order divisible by ℓ.*

Proof: Suppose 1 holds. Then σ is twist-equivalent to a representation of the form

$$\begin{pmatrix} \psi & u \\ 0 & 1 \end{pmatrix}$$

for some character ψ where u is a cocycle representing a class

$$x \in H^1(G_p, k(\psi)).$$

Since $\sigma|_{I_p}$ is indecomposable, the image of x in $H^1(I_p, k(\psi))^{G_p}$ does not vanish. This group is isomorphic to $\mathrm{Hom}_{G_p}(\mu_\ell(\bar{\mathbf{Q}}_p), k(\psi))$, which vanishes unless $\psi = \chi$. Moreover if χ is non-trivial, then $H^1(G_p, k(1))$ is one-dimensional over k, so 2 follows.

The implications $2 \Rightarrow 3 \Rightarrow 4$ are clear. If 4 holds, then 1 follows from the fact that $\sigma(G_p)$ is contained in the normalizer of the ℓ-Sylow subgroup of $\sigma(I_p)$.

Proposition 2.3. *The following are equivalent:*

1. *σ is irreducible and $\sigma|_{I_p}$ is reducible.*
2. *σ is equivalent to a representation of the form $\mathrm{Ind}_{G_M}^{G_p} \xi$, where M is the unramified quadratic extension of $\mathbf{Q}_p$ and ξ is a character of G_M not equal to its conjugate under the action of $\mathrm{Gal}(M/\mathbf{Q}_p)$.*
3. *$\tilde{\sigma}(I_p)$ is cyclic of order not divisible by ℓ, and $\tilde{\sigma}(G_p)$ is dihedral of twice that order.*
4. *$\tilde{\sigma}(I_p)$ is cyclic of order not divisible by ℓ, and $\tilde{\sigma}(G_p)$ is not cyclic.*

Proof: Suppose that 1 holds. Consider the action of G_p on $\mathbf{P}^1(k)$ gotten from $\tilde{\sigma}$. Note first that $\tilde{\sigma}(I_p)$ is nontrivial. Let S denote the set of elements in $\mathbf{P}^1(k)$ fixed by I_p. Since $\sigma|_{I_p}$ is reducible, S is not empty. Since σ is irreducible, S has no elements fixed by G_p and it follows that S has exactly two elements. Moreover G_p acts transitively on S via the unramified quadratic character, so 2 holds.

Suppose next that 2 holds. Then $\tilde{\sigma}(G_p)$ is a dihedral group in which $\tilde{\sigma}(G_M)$ is a cyclic subgroup of index two and order not divisible by ℓ. Since

$M^\times = \mathbf{Q}_p^\times \mathcal{O}_M^\times$, we see from local class field theory that

$$\xi^{-1}\xi'(G_M) = \xi^{-1}\xi'(I_p),$$

where ξ' is the conjugate of ξ, and 3 follows. The implication $3 \Rightarrow 4$ is clear, and $4 \Rightarrow 1$ follows from the converse of the corresponding one in Proposition 2.1.

Proposition 2.4. *The following are equivalent:*

1. $\sigma|_{I_p}$ *is irreducible.*
2. p *is odd and* σ *is equivalent to a representation of the form* $\mathrm{Ind}_{G_M}^{G_p} \xi$, *where* M *is a ramified quadratic extension of* $\mathbf{Q}_p$ *and* ξ *is a character of* G_M *whose restriction to* I_M *is not equal to its conjugate under the action of* $\mathrm{Gal}(M/\mathbf{Q}_p)$, *or* $p = 2$ *and the restriction of* σ *to the wild inertia subgroup of* G_p *is irreducible.*
3. $\tilde{\sigma}(I_p)$ *is dihedral of order* $2p^r$ *for some* $r \geq 1$ *and* $\tilde{\sigma}(G_p)$ *is dihedral of order dividing* $4p^r$, *or* $p = 2$, $\tilde{\sigma}(I_p)$ *(respectively* $\tilde{\sigma}(G_p)$*) is isomorphic to* D_4 *(respectively* A_4*),* A_4 *(respectively* A_4*) or* A_4 *(respectively* S_4*).*
4. $\tilde{\sigma}(I_p)$ *is not cyclic.*

Proof: Suppose that 1 holds and furthermore that $\sigma|_{P_p}$ is irreducible, where P_p is the wild inertia subgroup of I_p. Consider the action of G_p on $\mathbf{P}^1(k)$ gotten from $\tilde{\sigma}$. Since $\tilde{\sigma}(I_p)$ is not cyclic, we see that $\tilde{\sigma}(P_p)$ is nontrivial. Let S denote the set of elements in $\mathbf{P}^1(k)$ fixed by P_p. Then S is not empty and has no elements fixed by I_p. It follows that S has exactly two elements and that I_p acts transitively. Therefore p is odd and G_p acts transitively on S via a ramified quadratic character. We deduce that 2 holds, where M is the corresponding quadratic extension of $\mathbf{Q}_p$. (We have that $\xi \neq \xi'$ on P_p, hence on I_M.)

Suppose now that 2 holds. First consider the case of odd p. Then $\tilde{\sigma}(G_p)$ (respectively, $\tilde{\sigma}(I_p)$) is dihedral, and $\tilde{\sigma}(G_M)$ (respectively, $\tilde{\sigma}(I_M)$) is a cyclic subgroup of index two. Letting U denote the kernel of the reduction map on $\mathcal{O}_M^\times$, we have $\mathbf{Q}_p^\times U = \mathbf{Q}_p^\times \mathcal{O}_M^\times$ has index two in $M^\times$. From local class field theory it follows that $\tilde{\sigma}(I_M) = \tilde{\sigma}(P_M)$ has p-power order and index at most two in $\tilde{\sigma}(G_M)$. We conclude that 3 holds.

In the case of $p = 2$, we see that $D = \tilde{\sigma}(P_p)$ is dihedral, since it is not cyclic and is a finite subgroup of $PGL_2(k)$ of 2-power order. Furthermore $\tilde{\sigma}(G_p)$ is contained in the normalizer of D. If D has order greater than 4, the normalizer is dihedral and we may use the same argument as in the case of odd p. If D has order 4, then the normalizer is isomorphic to S_4, and 4 follows.

The implication $3 \Rightarrow 4$ is clear, as is $4 \Rightarrow 1$ (in view of the preceding propositions).

3. Minimally ramified liftings

For a fixed representation

$$\sigma : G_p \to \mathrm{GL}_2(k),$$

we consider liftings

$$\theta : G_{\mathbf{Q}} \to \mathrm{GL}_2(R)$$

of σ, where R is a complete local Noetherian $W(k)$-algebra with residue field k. We shall now say what it means for θ to be minimally ramified. We use $\tilde{}$ to denote composition with the Teichmüller lift

$$k^\times \to W(k)^\times \to R^\times.$$

Definition 3.1.

1. If σ is of type **P** or **V** , then

$$\sigma|_{I_p} \sim \begin{pmatrix} \xi_1 & 0 \\ 0 & \xi_2 \end{pmatrix},$$

and we say θ is *minimally ramified* if

$$\theta|_{I_p} \sim \begin{pmatrix} \tilde{\xi}_1 & 0 \\ 0 & \tilde{\xi}_2 \end{pmatrix}.$$

2. If σ is of type **S** , then

$$\sigma|_{I_p} \sim \xi \otimes \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix},$$

and we say θ is *minimally ramified* if

$$\theta|_{I_p} \sim \tilde{\xi} \otimes \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}.$$

3. If σ is of type **H** , then we say θ is *minimally ramified* if $\det \theta|_{I_p}$ is the Teichmüller lift of $\det \sigma|_{I_p}$.

Remark 3.2. First note that if χ is a character of $G_p \to k^\times$, then θ is a minimally ramified lifting of σ if and only if $\tilde{\chi} \otimes \theta$ is a minimally ramified lifting of $\chi \otimes \sigma$.

Remark 3.3. If σ is of type **P** , then it has a twist which is either unramified or of type B in the terminology of [W]. Note that if σ is unramified, then θ is minimally ramified if and only if θ is unramified.

Remark 3.4. If σ is of type **S** , then it has a twist of type A in the terminology of [W]. Recall that if θ arises from the ℓ-adic Tate module of an elliptic curve E over $\mathbf{Q}_p$ with split multiplicative reduction, then $\theta|_{I_p}$ is equivalent to a representation of the form $\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$. It is minimally ramified if and only if σ is ramified if and only if $v_p(\Delta_E)$ is divisible by ℓ.

Remark 3.5. Suppose now that σ is type **V** . If $p \not\equiv -1 \bmod \ell$, then σ is type C in the terminology of [W]. Suppose instead that $p \equiv -1 \bmod \ell$ and write $\sigma = \mathrm{Ind}_{G_M}^{G_p} \xi$ as in proposition 2.3. Let $\mu : G_M \to \mathcal{O}^\times$ be a ramified character of G_M of ℓ-power order. Then

$$\theta = \mathrm{Ind}_{G_M}^{G_p} \bar{\xi}\mu \tag{3}$$

is a lifting of σ which is *not* minimally ramified.

Remark 3.6. Now consider σ of type **H** . Suppose that

$$\sigma|_{I_p} \sim \mathrm{Ind}_{I_M}^{I_p} \xi$$

as in proposition 2.4. Then θ is minimally ramified if and only if

$$\theta|_{I_p} \sim \mathrm{Ind}_{I_M}^{I_p} \bar{\xi}.$$

Remark 3.7. Suppose that $\theta : G_p \to \mathrm{GL}_2(\mathcal{O})$ is a minimally ramified lifting of σ, where $\mathcal{O}$ is the ring of integers of a finite extension of the field of fractions of $W(k)$. Then $\det\theta|_{I_p}$ is the Teichmüller lift of $\det\sigma|_{I_p}$ and the Artin conductors of θ and σ coincide. In [W] and [TW] a technical hypothesis is imposed to ensure that a partial converse holds. This hypothesis rules out the existence of liftings as in (3) and facilitates the characterization of the modular forms which give rise to minimally ramified liftings. The main contribution of [D] is to dispense with that hypothesis.

4. UNIVERSAL DEFORMATION RINGS

Now consider an irreducible representation

$$\bar{\rho} : G_{\mathbf{Q}} \to \mathrm{GL}_2(k).$$

For each prime p we fix an embedding of $\bar{\mathbf{Q}}$ in $\bar{\mathbf{Q}}_p$ and regard G_p as a decomposition group in $G_{\mathbf{Q}}$. We suppose that $\bar{\rho}|_{G_\ell}$ is semistable in the sense of [DDT], section 2.4.

Suppose that K is a finite extension of the field of fractions of $W(k)$. Let $\mathcal{O}$ denote the integral closure of $W(k)$ in K; thus $\mathcal{O}$ is a complete discrete valuation ring with residue field k. We consider liftings of $\bar{\rho}$ of the form

$$\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(R),$$

where R is in the category $\mathbf{C}$ of local complete $\mathcal{O}$-algebras with residue field k. A deformation of $\bar{\rho}$ is an isomorphism class of such liftings (see [dSL] (2.1), (2.2)).

If Σ is a finite set of primes, we say that ρ is type Σ if

1. $\chi_\ell^{-1} \det\rho$ has finite order not divisible by ℓ;
2. ρ is minimally ramified outside Σ;
3. ρ is semistable at ℓ in the sense of [DDT].

The notion depends only on the isomorphism class of ρ and is independent of the choice of embeddings of $\bar{\mathbf{Q}}$ in $\bar{\mathbf{Q}}_p$.

Consider the functor which associates to R the set of deformations of $\bar{\rho}$ of type Σ. The type Σ restriction satisfies the conditions listed at the beginning of §6 of [dSL] (see also §29 of [M] and §2.4 of [DDT]). From [dSL] (2.4) and (6.1) we conclude that the functor is represented by a complete local $\mathcal{O}$-algebra R_Σ, the identity map of R_Σ corresponding to the universal deformation of type Σ:

$$\rho_\Sigma^{\text{univ}} : G_{\mathbf{Q}} \to \mathrm{GL}_2(R_\Sigma).$$

Suppose now that we are given a lifting

$$\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathcal{O})$$

of type Σ. The universal property of R_Σ yields a surjective morphism

$$\pi : R_\Sigma \to \mathcal{O}$$

such that ρ is equivalent to the pushforward of $\rho_\Sigma^{\text{univ}}$. Let $\mathfrak{p}$ denote the kernel of π. We define the group

$$H^1_\Sigma(G_{\mathbf{Q}}, (\mathrm{ad}^0\rho) \otimes (K/\mathcal{O}))$$

as in §2.7 of [DDT]. A generalization of results of Mazur (see §23–25 of [M]) yields a canonical isomorphism

$$\mathrm{Hom}_{\mathcal{O}}(\mathfrak{p}/\mathfrak{p}^2, K/\mathcal{O}) \cong H^1_\Sigma(G_{\mathbf{Q}}, (\mathrm{ad}^0\rho) \otimes_{\mathcal{O}} (K/\mathcal{O})). \tag{4}$$

5. HECKE ALGEBRAS

Recall that given a newform

$$f(\tau) = \sum a_n(f) e^{2\pi i n \tau}$$

of weight 2, level N_f and character ψ_f, a construction of Eichler and Shimura (see [Ro]) associates to f a continuous representation

$$\rho_f : G_{\mathbf{Q}} \to \mathrm{GL}_2(\bar{\mathbf{Q}}_\ell),$$

where we have fixed embeddings $\bar{\mathbf{Q}} \to \mathbf{C}$ and $\bar{\mathbf{Q}} \to \bar{\mathbf{Q}}_\ell$. The representation ρ_f is characterized up to isomorphism by the following property: For all primes p not dividing $N_f\ell$, ρ_f is unramified at p and the characteristic polynomial of $\rho_f(\mathrm{Frob}_p)$ is

$$X^2 - a_p(f)X + \psi_f(p)p.$$

We wish to continue working over the ring $\mathcal{O}$ introduced above, so we also fix an embedding $\bar{\mathbf{Q}}_\ell \to \bar{K}$ and view ρ_f as taking values in $\mathrm{GL}_2(K_f)$, where K_f is the subfield of $\bar{K}$ generated by K and the Fourier coefficents of f. We denote the ring of integers $\mathcal{O}_f$, which we regard as an object of $\mathbf{C}$. Define

$$\bar{\rho}_f : G_{\mathbf{Q}} \to \mathrm{GL}_2(k)$$

as the semisimplification of the reduction of ρ_f.

We assume that our fixed representation $\bar{\rho}$ is modular in the sense that it is isomorphic to $\bar{\rho}_f$ for some weight 2 newform f. We let Φ_Σ denote the set of newforms g such that ρ_g is a deformation of $\bar{\rho}$ of type Σ and N_g is not divisible by ℓ^2.

Theorem 5.1. *If $\bar{\rho}$ is modular, then $\Phi_\emptyset \neq \emptyset$.*

This is a refinement of Serre's ϵ-conjecture for which a crucial ingredient is Ribet's theorem [Ril] (see [E]). The result stated here is a consequence of [D] which builds on the work Ribet and many others.

For each g in Φ_Σ, we consider the map $R_\Sigma \to \mathcal{O}_g$ corresponding to ρ_g. We then define

$$\mathbf{T}_\Sigma \subset \prod_{g \in \Phi_\Sigma} \mathcal{O}_g$$

as the image of R_Σ. Since R_Σ is topologically generated by traces, we may also regard $\mathbf{T}_\Sigma$ as the $\mathcal{O}$-subalgebra generated by the elements

$$T_p = (a_p(g))_{g \in \Phi_\Sigma}$$

for primes p not dividing $N\ell$. We wish to prove that the surjective map

$$\phi_\Sigma : R_\Sigma \to \mathbf{T}_\Sigma$$

is an isomorphism. Note that Φ_Σ gives rise to a type Σ deformation

$$\rho_\Sigma^{\mathrm{mod}} : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{T}_\Sigma)$$

of $\bar{\rho}$, such that for each $g \in \Phi$, the composition with the projection to $\mathrm{GL}_2(\mathcal{O}_g)$ is equivalent to ρ_g.

For finite sets of primes $\Sigma \supset \Theta$, there is a natural surjective homomorphism $R_\Sigma \to R_\Theta$ defined by regarding $\rho_\Theta^{\mathrm{univ}}$ as a deformation of $\bar{\rho}$ of type Σ. We have also the natural surjection $\mathbf{T}_\Sigma \to \mathbf{T}_\Theta$ so that the diagram

$$\begin{array}{ccc} R_\Sigma & \stackrel{\phi_\Sigma}{\to} & \mathbf{T}_\Sigma \\ \downarrow & & \downarrow \\ R_\Theta & \stackrel{\phi_\Theta}{\to} & \mathbf{T}_\Theta \end{array} \tag{5}$$

commutes.

6. THE MAIN RESULTS

Recall our assumption that ℓ is odd and $\bar{\rho}$ is semistable at ℓ. We let $L = \mathbf{Q}(\sqrt{\varepsilon\ell})$, where $\varepsilon = (-1)^{(\ell-1)/2}$. We suppose that Σ is an arbitrary finite set of primes. The main result is the following:

Theorem 6.1. *If $\bar{\rho}|_{G_L}$ is irreducible, then ϕ_Σ is an isomorphism and $\mathbf{T}_\Sigma$ is a complete intersection.*

We shall sketch the proof below referring to [D] for the full details.

Corollary 6.2. *Suppose that $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathcal{O})$ is continuous and unramified outside a finite set of primes. Suppose that ρ is semistable at ℓ and $\bar{\rho}|_{G_L}$ is irreducible. If $\bar{\rho}$ is modular, then ρ is modular.*

Applying the Langlands-Tunnell theorem (see [G]) as in [Ru], we conclude:

Corollary 6.3. *Suppose that E is an elliptic curve over $\mathbf{Q}$ with good or multiplicative reduction at 3, and that $[\mathbf{Q}(E[3]) : \mathbf{Q}] = 16$ or 48. Then E is modular.*

We refer to [Ru] for the deduction of theorem 1.1 from corollaries 6.2 and 6.3.

7. Sketch of proof

7.1. **Vague principle.** A formulation of the problem such as theorem 6.1 enables us to use tools from commutative algebra. We shall use information about the vertical maps in (5) and one of the horizontal maps to prove that the other horizontal map is an isomorphism. The information about $R_\Sigma \to R_\Theta$ comes from the description of tangent spaces in terms of Galois cohomology (4); the information about $\mathbf{T}_\Sigma \to \mathbf{T}_\Theta$ comes from the connection with congruences between modular forms.

7.2. **Some preparation.** We begin with two reduction steps and a definition.

One can check that if ψ is a character $G_{\mathbf{Q}} \to k^\times$ unramified outside ℓ, then theorem 6.1 holds for $\bar{\rho}$ if and only if it holds for $\bar{\rho}' = \bar{\rho} \otimes \chi$. Indeed if we define $\phi'_\Sigma : R'_\Sigma \to \mathbf{T}'_\Sigma$ using $\bar{\rho}'$ instead of $\bar{\rho}$, then we obtain a natural commutative diagram

$$\begin{array}{ccc} R_\Sigma & \xrightarrow{\sim} & R'_\Sigma \\ \downarrow{\scriptstyle \phi_\Sigma} & & \downarrow{\scriptstyle \phi'_\Sigma} \\ \mathbf{T}_\Sigma & \xrightarrow{\sim} & \mathbf{T}'_\Sigma. \end{array}$$

We can therefore assume that for each prime $p \neq \ell$ such that $\bar{\rho}|_{G_p}$ is reducible (i.e., $\mathbf{P}$ or $\mathbf{S}$), we have $\bar{\rho}^{I_p} \neq 0$.

We also find that theorem 6.1 is well-behaved under extension of scalars. More precisely, suppose that K' is a finite extension of K. Defining

$$\phi'_\Sigma : R'_\Sigma \to \mathbf{T}'_\Sigma$$

using K' instead of K, we find that there is a natural commutative diagram

$$\begin{array}{ccc} R_\Sigma \otimes_{\mathcal{O}} \mathcal{O}' & \xrightarrow{\sim} & R'_\Sigma \\ \downarrow{\scriptstyle \phi_\Sigma \otimes 1} & & \downarrow{\scriptstyle \phi'_\Sigma} \\ \mathbf{T}_\Sigma \otimes_{\mathcal{O}} \mathcal{O}' & \xrightarrow{\sim} & \mathbf{T}'_\Sigma, \end{array}$$

where $\mathcal{O}'$ is the ring of integers of K'. One deduces from this that if theorem 6.1 holds for some K, then it holds for all K. In particular, we may assume that there is an $\mathcal{O}$-algebra homomorphism $\mathbf{T}_\emptyset \to \mathcal{O}$.

In view of remark 3.7, we must exercise extra care with primes p such that $\bar{\rho}|_{G_p}$ is of type $\mathbf{V}$. We denote by P the set of such vexing primes. (In [W] and [TW], it is assumed that P consists only of primes which are not congruent to -1 mod ℓ.)

7.3. **The case $\Sigma = \emptyset$.** Recall that the strategy of Wiles and Taylor-Wiles in the "minimal case" is to choose, for each $n \geq 1$, a certain set $Q = Q_n$ consisting of primes congruent to 1 mod ℓ^n. These sets Q are chosen so that R_Q and $R_\emptyset$ can be topologically generated as an $\mathcal{O}$-algebra by r elements, where r is the cardinality of Q. Moreover the choice is made so that $\mathbf{T}_Q$ and $\mathbf{T}_\emptyset$ can be related using their natural structure as algebras over a group ring where the group is generated by r elements. One then proves $\phi_\emptyset$ is an isomorphism using the arguments of §3 of [TW] and Chapter 3 of [W], or using the Taylor-Wiles-Faltings criterion ([TW], Appendix or [DDT], §3.4). Alternatively, using Rubin's simplification of the isomorphism criterion (see [dSRS]), it suffices to choose a single set $Q = Q_n$ as in [TW], where n is made explicit.

Our strategy is the same, but the set P introduces several complications. A minor complication is that we use a version over $\mathcal{O}$ of the isomorphism criterion (see §5 of [D]). We shall now state such a version along the lines of Rubin's simplification, leaving it as an exercise to make the necessary modifications to the proof of Criterion II of [dSRS].

We fix an integer $r \geq 0$ and consider power series rings

$$\mathcal{O}[[S]] = \mathcal{O}[[S_1, \ldots, S_r]] \qquad \text{and} \qquad \mathcal{O}[[X]] = \mathcal{O}[[X_1, \ldots, X_r]].$$

Let $\mathfrak{m}$ denote the maximal ideal of $\mathcal{O}[[S]]$. Recall that the polynomial

$$f(x) = \prod_{i=0}^{r} (x + i)$$

satisfies $f(n)/(r+1)! = \text{length}_{\mathcal{O}}(\mathcal{O}[[S]]/\mathfrak{m}^n)$ for all integers $n \geq 1$. We also fix $\mathcal{O}$-algebra homomorphisms

$$\mathcal{O}[[S]] \to \mathcal{O}[[X]] \to R \to T \tag{6}$$

with $\mathcal{O}[[X]] \to R$ and $R \to T$ surjective. Suppose that $T/(S_1, \ldots, S_r)T$ is finitely generated as an $\mathcal{O}$-module; let s denote its rank and t the $\mathcal{O}$-length of its torsion.

Theorem 7.1. *Suppose that there are positive integers d and N such that*

1. $d \geq st + s + t$,
2. $f(N) + f(dN - d) - f(dN) > 0$,
3. $\mathcal{O}[[S]]/\mathfrak{m}^N \to T/\mathfrak{m}^N T$ *is injective.*

Then

- $R/(S_1,\dots,S_r)R \to T/(S_1,\dots,S_r)T$ *is an isomorphism,*
- $T/(S_1,\dots,S_r)T$ *is a local complete intersection,*
- $s > 0$ *and* $t = 0$.

We shall apply the criterion with $R = R_Q$ and $\mathbf{T} = \mathbf{T}_Q$ for a certain set Q as in [TW]. We shall explain below how r, d, N and Q are to be chosen, and the maps in (6) are to be defined.

For arbitrary Σ, let I_Σ denote the kernel of the map $R_\Sigma \to R_\emptyset$. One can check that the kernel of the natural surjection

$$\mathbf{T}_\Sigma/I_\Sigma\mathbf{T}_\Sigma \to \mathbf{T}_\emptyset$$

is torsion. In particular, the rank of $\mathbf{T}_\Sigma/I_\Sigma\mathbf{T}_\Sigma$ is independent of Σ, and we denote it s'. We denote by t' the $\mathcal{O}$-length of the torsion submodule of $\mathbf{T}_P/I_P\mathbf{T}_P$. We set $d = s't' + s' + t'$, $r = \dim_k H^1_\emptyset(G_\mathbf{Q}, \mathrm{ad}^0\bar\rho(1))$ and choose N so that the inequality of theorem 7.12 is satisfied. (Note that $f(N) + f(dN - d) - f(dN)$ is a polynomial with leading term N^{r+1}.)

By the same Galois cohomology argument as in §4 of [TW] (or see [dSh] or [DDT]), we choose a finite set of primes Q such that

- $\#Q = r$,
- R_Q can be topologically generated as an $\mathcal{O}$-algebra by r elements,
- if $q \in Q$, then the following hold:
  - $q \equiv 1 \bmod \ell^N$;
  - $\bar\rho$ is unramified at q;
  - $\bar\rho(\mathrm{Frob}_q)$ has distinct eigenvalues.

Since R_Q is generated by r elements as an $\mathcal{O}$-algebra, we can define a surjective homomorphism $\mathcal{O}[[X]] \to R_Q$. Let G denote the maximal quotient of $\prod_{q\in Q}(\mathbf{Z}/q\mathbf{Z})^\times$ of ℓ-power order. We endow $R_{P\cup Q}$, hence R_Q, with the structure of an $\mathcal{O}[G]$-algebra as in [TW], appendix (or see [dSh] or [DDT]). Choosing generators $g_1,\dots,g_r$ for g, we define a surjection

$$\begin{array}{ccc} \mathcal{O}[[S]] & \to & \mathcal{O}[G] \\ S_i & \mapsto & g_i - 1 \end{array}$$

whose kernel is contained in $\mathfrak{m}^N$. We then define the $\mathcal{O}$-algebra homomorphism $\mathcal{O}[[S]] \to \mathcal{O}[[X]]$ so that the diagram

$$\begin{array}{ccc} \mathcal{O}[[S]] & \longrightarrow & \mathcal{O}[[X]] \\ \downarrow & & \downarrow \\ \mathcal{O}[G] & \longrightarrow & R_Q \end{array}$$

commutes.

The verification of hypothesis 3 can be viewed as the main obstacle in improving the methods of [TW] and [W] to cover the setting of theorem 6.1. Recall that Taylor and Wiles use a method of de Shalit to prove that (under

their hypotheses), $\mathbf{T}_Q$ is free over $\mathcal{O}[G]$ and $\mathbf{T}_Q/I_Q\mathbf{T}_Q \xrightarrow{\sim} \mathbf{T}_\emptyset$ (see §2 of [TW] or [dSh], or see §4.3 of [DDT] for an alternative argument using the q-expansion principle). The key observation made in [D] is that it suffices to prove the following:

Lemma 7.2. *There exists a nonzero $\mathbf{T}_Q$-module which is free over $\mathcal{O}[G]$.*

The proof of the lemma is very technical and is related to the methods of [DT]. We refer the reader to §4 of [D] for details, mentioning here only that it relies on the Jacquet-Langlands correspondence and a cohomological construction. We also point out that to prove the lemma and other results used below on the fine structure of the algebras $\mathbf{T}_\Sigma$, one first realizes them as completions of Hecke algebras acting on spaces of modular forms. (See for example §4.1 and §4.2 of [DDT].)

To verify that the hypothesis 1 of the theorem is satisfied, one uses that $\mathcal{O}[G] \to R_Q$ was defined so that the augmentation ideal of $\mathcal{O}[G]$ maps onto $I_Q = \ker(R_Q \to R_\emptyset)$. Thus we have

$$s = \operatorname{rank}_{\mathcal{O}}(\mathbf{T}_Q/I_Q\mathbf{T}_Q) = \operatorname{rank}_{\mathcal{O}}\mathbf{T}_\emptyset = \operatorname{rank}_{\mathcal{O}}(\mathbf{T}_P/I_Q\mathbf{T}_P) = s'.$$

The arguments of [TW] discussed (or [DDT] §4.3) can be used to show that the natural map

$$\mathbf{T}_{P\cup Q}/(S_1, \ldots, S_r)\mathbf{T}_{P\cup Q} \to \mathbf{T}_P \tag{7}$$

is an isomorphism (see [D], lemma 3.3). One then deduces that

$$\mathbf{T}_{P\cup Q}/I_{P\cup Q}\mathbf{T}_{P\cup Q} \xrightarrow{\sim} \mathbf{T}_P/I_P,$$

from which it follows that $t \le t'$ and $d \le d'$.

We now apply theorem 7.1 to conclude that

$$R_Q/I_Q \to \mathbf{T}_Q/I_Q\mathbf{T}_Q$$

is an isomorphism, and these rings are complete intersections and torsion-free over $\mathcal{O}$. From this follows theorem 6.1 in the case $\Sigma = \emptyset$.

7.4. **The case of arbitrary Σ.** Our situation now is that we have a commutative diagram of surjective $\mathcal{O}$-algebra homomorphisms

$$\begin{array}{ccc} R_\Sigma & \xrightarrow{\phi_\Sigma} & \mathbf{T}_\Sigma \\ \downarrow & & \downarrow \\ R_\emptyset & \xrightarrow{\phi_\emptyset} & \mathbf{T}_\emptyset; \end{array}$$

we know that the bottom row is an isomorphism and the rings are local complete intersections, and we wish to prove this holds for the top row.

Recall that we have assumed the existence of a map $\mathbf{T}_\emptyset \to \mathcal{O}$ of $\mathcal{O}$-algebras. Such a homomorphism necessarily corresponds to newform f with coefficients in $\mathcal{O}$ such that ρ_f is a deformation of $\bar{\rho}$ of type $\emptyset$.

For arbitrary Θ, we write $\mathfrak{p}_\Theta$ for the kernel of $R_\Theta \to \mathcal{O}$, and $\mathfrak{P}_\Theta$ for the kernel of $\pi_\Theta : \mathbf{T}_\Theta \to \mathcal{O}$. We consider the $\mathcal{O}$-module $\Phi_\Theta = \mathfrak{p}_\Theta/\mathfrak{p}_\Theta^2$ and the $\mathcal{O}$-ideal $\eta_\Theta = \pi_\Theta(\mathrm{Ann}_{\mathbf{T}_\Theta}\mathfrak{P}_\Theta)$. We omit the subscript Θ when $\Theta = \emptyset$. According to the Wiles-Lenstra criterion, Criterion I of [dSRS], we know that

$$\mathrm{length}_{\mathcal{O}}(\Phi) = \mathrm{length}_{\mathcal{O}}(\mathcal{O}/\eta),$$

and we wish to prove that

$$\mathrm{length}_{\mathcal{O}}(\Phi_\Sigma) \le \mathrm{length}_{\mathcal{O}}(\mathcal{O}/\eta_\Sigma).$$

Using (4), one obtains as in §4.2 of [Ri2]

$$\mathrm{length}_{\mathcal{O}}(\Phi_\Sigma) \le \mathrm{length}_{\mathcal{O}}(\Phi) + \sum_{p\in\Sigma} d_p, \tag{8}$$

where d_p is the length of

- $H^0(G_p, \mathrm{ad}^0\rho_f \otimes_{\mathcal{O}} K/\mathcal{O}(1))$ if $p \neq \ell$;
- $\mathcal{O}/(a_\ell(f)^2 - \psi_f(\ell))$ if $p = \ell$ does not divide N_f;
- 0 otherwise.

(We have used here that ρ_f is of type $\emptyset$.)

Using that $\mathbf{T}_\Sigma$ is Gorenstein for $\Sigma \supset P$ (Wiles' generalization of results of Mazur and others discussed in [Ti]), together with Wiles' calculations of the change in η discussed in §4.3 of [Ri2], we find that

$$\mathrm{length}_{\mathcal{O}}(\mathcal{O}/\eta_\Sigma) \ge \mathrm{length}_{\mathcal{O}}(\mathcal{O}/\eta_P) + \sum_{p\in\Sigma-P} d_p, \tag{9}$$

(provided $\Sigma \supset P$). We complement this with the inequality

$$\mathrm{length}_{\mathcal{O}}(\mathcal{O}/\eta_P) \ge \mathrm{length}_{\mathcal{O}}(\mathcal{O}/\eta) + \sum_{p\in P} d_p \tag{10}$$

established by lemma 3.6 of [D].

Applying the Wiles-Lenstra criterion together with (8), (9) and (10), we conclude that Φ_Σ is an isomorphism if $\Sigma \supset P$. (This is all that is proved in [D] and all that is needed for the corollaries.) We leave it as an exercise for the reader to treat the case of arbitrary Σ by showing that if $P_0 \subset P \subset \Sigma$, then $\prod_{p\in P_0}(p+1)$ is an element of $\pi_\Sigma(J)$, where J is the annihilator in $\mathbf{T}_\Sigma$ of the kernel of $\mathbf{T}_\Sigma \to \mathbf{T}_{\Sigma-P_0}$.

Acknowledgement. The author was supported by the United Kingdom's EPSRC (#GR/J4761) while this paper was written.

REFERENCES

[DDT] H. Darmon, F. Diamond, R. Taylor, *Fermat's Last Theorem*, in Current Developments in Mathematics, 1995, International Press, 1–154.

[dSh] E. de Shalit, *Hecke rings and universal deformation rings*, this volume.

[dSL] B. de Smit, H. Lenstra, *Explicit construction of unversal deformation rings*, this volume.

[dSRS] B. de Smit, K. Rubin, R. Schoof, *Criteria for complete intersections*, this volume.

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

[Ri2] F. Diamond, K. Ribet, *p-adic modular deformations and Wiles' "Main Conjecture,"* this volume.

[DT] F. Diamond, R. Taylor, *Lifting modular mod l representations*, Duke Math J. **74** (1994), 253–269.

[E] B. Edixhoven, *Serre's conjecture*, this volume.

[G] S. Gelbart, *Three lectures on the modularity of $\bar{\rho}_{E,3}$ and the Langlands reciprocity conjecture*, this volume.

[M] B. Mazur, *An introduction to the deformation theory of Galois representations*, this volume.

[Ri1] K. Ribet, *On modular representations of* $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ *arising from modular forms*, Inv. Math. **100** (1990), 431–476.

[Ro] D. Rohrlich, *Modular functions and modular curves*, this volume.

[Ru] K. Rubin, *Modularity of mod 5 representations*, this volume.

[S] J.H. Silverman, *A survey of the arithmetic theory of elliptic curves*, this volume.

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

[Ti] J. Tilouine, *Hecke algebras and the Gorenstein property*, this volume.

[W] A. Wiles, *Modular elliptic curves and Fermat's Last Theorem*, Annals of Math. **141** (1995), 443–551.