Non-minimal modularity lifting in weight one

By *Frank Calegari* at Chicago

Abstract. We prove an integral $R = \mathbf{T}$ theorem for odd two-dimensional p-adic representations of $G_{\mathbf{Q}}$ which are unramified at p, extending results of [5] to the non-minimal case. We prove, for any p, the existence of Katz modular forms modulo p of weight one which do not lift to characteristic zero.

1. Introduction

The main innovation of [5] was to develop a framework for modularity lifting theorems in contexts in which the Taylor–Wiles method did not apply. One of the main examples in [5, Theorem 1.4] was a minimal modularity lifting theorem for odd two-dimensional Galois representations which are unramified at p. This result was simultaneously a generalization and a specialization of the main theorem of Buzzard–Taylor [4]; generalized because it related Galois representations modulo ϖ^n to Katz modular forms of weight one modulo ϖ^n neither of which need lift to characteristic zero, and specialized because it required a minimality hypothesis at primes away from p. One of the goals of the present paper is to provide a theorem which is a new proof of many cases of [4] in the spirit of [5]. Our methods could be viewed as hybrid of both [5] and [4] in the following sense: as in [5], we prove an integral $R = \mathbf{T}$ theorem for torsion representations by working directly in weight one, however, as a crucial input, we use ordinary modularity lifting theorems in higher weight (as in [4], although we only need to work in weight p) in order to show that the patched Hecke modules see every component of the generic fibre of the global deformation ring. In order to simplify some of our arguments, we do not strive for maximal generality. The assumption that the representations are unramified at p, however, seems essential for the method (if one does not use base change), in contrast to [2]. Let us fix a prime $p > 2$ and a local field $[E : \mathbf{Q}_p] < \infty$ with ring of integers $\mathcal{O}$ and residue field $k = \mathcal{O}/\varpi$.

Theorem 1.1. *Let $p > 2$, and let $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathcal{O})$ be a continuous odd Galois representation ramified at finitely many primes and unramified at p. Suppose that $\overline{\rho}$ is absolutely irreducible. If ρ is ramified at a prime ℓ, assume that $\rho|D_\ell$ is reducible. Then ρ is modular of weight one.*

This result will be deduced from our main result, which is an integral $R = \mathbf{T}$ theorem which we now describe. Let

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

be a continuous absolutely irreducible odd representation unramified at p. For each ℓ, let ϵ denote the cyclotomic character. Let ψ denote the Teichmüller lift of $\det(\bar{\rho})$. Let $N = S \cup P$ be a set of primes not containing p such that ψ is ramified exactly at the primes contained in P and unramified at primes contained S. By abuse of notation, we also let N denote the product of the conductor of ψ with the primes in S. We consider the functor D^1 from complete local Noetherian $\mathcal{O}$-algebras $(A, \mathfrak{m})$ with residue field k defined (informally) as follows. Fix a collection of elements $a_\ell \in k$ for ℓ dividing N. Let $D^1(A)$ consist of deformations ρ to A together with a collection of elements $\alpha_\ell \in A$ for $\ell \in N$ such that:

(i) $\det(\rho) = \psi$.

(ii) ρ is unramified outside $N = S \cup P$.

(iii) If $\ell \in P$, then

$$\rho|D_\ell \simeq \chi^{-1}\psi|_{D_\ell} \oplus \chi$$

for some unramified character χ with $\chi(\mathrm{Frob}_\ell) = \alpha_\ell \equiv a_\ell \bmod \mathfrak{m}$.

(iv) If $\ell \in S$, then

$$\rho|D_\ell \simeq \begin{pmatrix} \chi^{-1}\psi|_{D_\ell} & * \\ 0 & \chi \end{pmatrix}$$

for some unramified character χ with $\chi(\mathrm{Frob}_\ell) = \alpha_\ell \equiv a_\ell \bmod \mathfrak{m}$.

In fact, the actual definition of $D^1(A)$ needs to be somewhat modified (see Section 2.2 for precise definitions), but this description will be valid for rings of integers such as $\mathcal{O}$. Naturally enough, we also assume that $D^1(k)$ is non-empty, and that k is also large enough to contain the eigenvalues of every element in the image of $\bar{\rho}$. The elements $a_\ell \in k$ are determined by $\bar{\rho}$ for primes in P, but not necessarily for primes in S, because when $\bar{\rho}$ is unramified at ℓ, there is a choice of eigenvalue for the unramified line. Hence D^1 will not strictly be a Galois deformation ring; we refer to such rings (and we have several in this paper) as modified deformation rings because they depend not only on $\bar{\rho}$ but also on some auxiliary data. The functor D^1 is representable by a complete local $\mathcal{O}$-algebra R^1. The ring R^1 comes with elements $a_\ell \in k$ and $\alpha_\ell \in R^1$ for ℓ dividing N. For all other primes ℓ, define $a_\ell \in k$ to be $\mathrm{tr}(\bar{\rho}(\mathrm{Frob}_\ell))$, including when $\ell = p$.

We are now ready to state out main theorem. By abuse of notation, let N denote the conductor of ψ times the primes in S – it is divisible exactly by the primes in $N = S \cup P$. Let $X_H(N)$ denote the quotient of $X_1(N)$ by the Sylow p-subgroup of $(\mathbf{Z}/N\mathbf{Z})^\times$. After enlarging S if necessary, we may assume that the curve $X_H(N)$ is a fine moduli space. ($X_H(N)$ will be automatically a fine moduli space if $p > 3$, see Section 2.3.)

Theorem 1.2. *Let $p > 2$. Let $\mathbf{T} \subset \mathrm{End}_{\mathcal{O}} H^0(X_H(N), \omega_{E/\mathcal{O}})$ be the $\mathcal{O}$-algebra generated by Hecke endomorphisms. Let $\mathfrak{m}$ be the maximal ideal of $\mathbf{T}$ generated by the elements $\langle \ell \rangle - \psi(\ell)$ for $\ell \nmid N$ and $T_\ell - a_\ell$ for all ℓ. Then there is an isomorphism $R^1 \simeq \mathbf{T}_{\mathfrak{m}}$.*

1.1. Theorem 1.2 implies Theorem 1.1. Suppose that $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathcal{O})$ is a continuous Galois representation ramified at $\ell \neq p$ satisfying the conditions of Theorem 1.1. Then

after a global twist and enlarging $\mathcal{O}$ if necessary to contain a choice of elements a_ℓ for ℓ dividing N, ρ gives rise to an element of $D^1(\mathcal{O})$. The modularity of ρ then follows.

As an application of Theorem 1.2, we prove the following:

Theorem 1.3. *Let p be any prime. There exists a Katz modular form*

$$f \in H^0(X(\Gamma)_{\mathbf{F}_p}, \omega)$$

for some level Γ prime to p which does not lift to characteristic zero.

The original argument of Wiles [26, 27] for modularity theorems at non-minimal level was to use an induction argument and a certain numerical criterion involving complete intersections which were finite over $\mathcal{O}$. This does not seem to be obviously generalizable to weight one – although one still has access to forms of Ihara's Lemma, the Hecke rings $\mathbf{T}^1$ are no longer complete intersections in general, and are certainly not flat over $\mathcal{O}$. It remains open as to whether one can proceed using such an argument. Instead, we use modularity theorems in weight p in order to show $R_Q^1/\varpi \simeq \mathbf{T}_{Q,\mathfrak{m}}^1/\varpi$ for various sets of auxiliary primes Q, and we then use this information to show that the patched Hecke modules in weight one are "big enough." To pass between weight one and weight p we *crucially* rely on q-expansions. For this reason, the methods of this paper will probably not be generalizable beyond GL(2) (although they may have implications for Hilbert modular forms of partial weight one). Note that, in writing the paper [5], we tried to avoid the use of q-expansions as much as possible, whereas the philosophy of this paper is quite the opposite.

Remark 1.4. The methods of our paper may well be able to handle more precise local deformation conditions than those considered above. However, these assumptions considerably simplify some aspects of the arguments. We particularly shun Diamond's vexing primes, which did indeed cause considerable vexation in [5]. In fact, we try so hard to avoid them that we assume that $\overline{\rho}|D_\ell$ is reducible, when certainly some such representations – for example those with $\overline{\rho}|I_\ell$ irreducible – may well be amenable to our methods.

Acknowledgement. The debt this paper owes to [5] is clear, and the author thanks David Geraghty for many conversations. We thank Mark Kisin for the explaining a proof of Lemma 2.6, and we also thank Brian Conrad for a related proof of the same result in the context of rigid analytic geometry. We thank Toby Gee and Patrick Allen for several useful comments. We also thank Gabor Wiese for the original idea of proving modularity theorems in weight one by working in weight p.

2. Preliminaries

2.1. Local modified deformation rings. Let $\overline{\rho} : G_{\mathbf{Q}_\ell} \to \mathrm{GL}_2(k)$ be a representation, and let R_ℓ^{univ} denote the universal framed local deformation ring, and ρ^{univ} the universal local deformation. We assume in this entire section that $p > 2$. The calculations in this section will mostly be concerned with the case that $\ell \neq p$. Fix a lift of Frobenius $\phi \in G_{\mathbf{Q}_\ell}$, and choose an eigenvalue a_ℓ of $\overline{\rho}(\phi)$, which, after enlarging k if necessary, we may assume to lie in k.

We define the universal modified framed local deformation ring $\widetilde{R}_\ell^{\mathrm{univ}}$ to be the localization of the ring

$$R_\ell^{\mathrm{univ}}[\alpha_\ell]/(\alpha_\ell^2 - \alpha_\ell \mathrm{Tr}(\rho^{\mathrm{univ}}(\phi)) + \det(\rho^{\mathrm{univ}}(\phi)))$$

at $(\alpha_\ell - a_\ell)$. The quadratic polynomial satisfied by α_ℓ is the characteristic polynomial of Frobenius.

Lemma 2.1. *If $\overline{\rho}(\phi)$ has distinct eigenvalues, then $\widetilde{R}_\ell^{\mathrm{univ}} \simeq R_\ell^{\mathrm{univ}}$. If $\overline{\rho}(\phi)$ does not have distinct eigenvalues, then $\widetilde{R}_\ell^{\mathrm{univ}}$ is a finite flat extension of R_ℓ^{univ} of degree two.*

Proof. If $\overline{\rho}(\phi)$ has distinct eigenvalues, then the characteristic polynomial of Frobenius is separable over k. Since R_ℓ^{univ} is complete, the polynomial also splits over R_ℓ^{univ} by Hensel's Lemma, and the quadratic extension above is, (before localization), isomorphic to $R_\ell^{\mathrm{univ}} \oplus R_\ell^{\mathrm{univ}}$. Localizing at $(\alpha_\ell - a_\ell)$ picks out the factor on which we have the congruence $\alpha_\ell \equiv a_\ell \mod \mathfrak{m}$. If the eigenvalues of $\overline{\rho}(\phi)$ are both a_ℓ, then the quadratic extension is already local. $\square$

A modified local deformation ring will simply be a quotient of $\widetilde{R}_\ell^{\mathrm{univ}}$. Proposition 3.1.2 of [11] proves the existence of quotients $R_\ell^{\mathrm{univ},\psi,\tau}$ of R_ℓ^{univ} which are reduced, $\mathcal{O}$-flat, equidimension of dimension 4, and such that, for any finite extension F/E, a map

$$x : R^{\mathrm{univ}} \to F$$

factors through $R^{\mathrm{univ},\psi,\tau}$ if and only if the corresponding F representation V_x has determinant ψ and is of type τ. (For this section ψ may be any unramified character.) For our purposes, it will suffice to consider the trivial type τ, which corresponds to representations on which

$$\rho_x : G_{\mathbf{Q}_\ell} \to \mathrm{GL}(V_x) = \mathrm{GL}_2(F)$$

restricted to the inertial subgroup $I_\ell \subset G_{\mathbf{Q}_\ell}$ has unipotent (and so possibly trivial) image.

Lemma 2.2. *Suppose that $p > 2$ and $\ell \neq p$. Let τ denote the trivial type. There exists a quotient $R_\ell^{\mathrm{mod},\psi} := \widetilde{R}_\ell^{\mathrm{univ},\psi,\tau}$ of $\widetilde{R}_\ell^{\mathrm{univ}}$ which is reduced, $\mathcal{O}$-flat, equidimensional of dimension 4, and such that, for any finite extension F/E, a map*

$$x : \widetilde{R}_\ell^{\mathrm{univ}} \to F$$

factors through $R_\ell^{\mathrm{mod},\psi}$ if and only if the corresponding F representation V_x has determinant ψ, is ordinary, and has an unramified quotient on which the action of Frob_ℓ is by the image of α_ℓ.

The arguments are very similar to those already in the literature, but for want of a reference which covers this case exactly, we give the details.

Proof. Suppose that the eigenvalues of $\overline{\rho}(\phi)$ are distinct. Then, under the isomorphism $R_\ell^{\mathrm{univ}} \simeq \widetilde{R}_\ell^{\mathrm{univ}}$, we may take $\widetilde{R}_\ell^{\mathrm{univ},\psi,\tau} = R_\ell^{\mathrm{univ},\psi,\tau}$. Hence we may assume that the eigenvalues are the same. Any representation

$$\rho_x : G_{\mathbf{Q}_\ell} \to \mathrm{GL}_2(F)$$

for which the image of inertia has non-trivial unipotent image is, up to twist, and enlarging F if necessary, an extension of F by $F(1)$. In particular, the ratios of the eigenvalues of $\rho_x(\phi)$ must be equal to ℓ. Since we are assuming the eigenvalues of $\overline{\rho}(\phi)$ coincide, then,

if $\ell \not\equiv 1 \bmod p$, $R_\ell^{\mathrm{univ},\psi,\tau}$ has no such quotients, and will consist precisely the unramified locus. In this case, we may take $\widetilde{R}_\ell^{\mathrm{univ},\psi,\tau}$ to be the double cover corresponding to the unramified locus with a choice of Frobenius eigenvalue. Hence we may assume that $\ell \equiv 1 \bmod p$, and in particular $\ell \not\equiv -1 \bmod p$.

Assume that $\overline{\rho}$ is unramified. The ring $R_\ell^{\mathrm{univ},\psi,\tau}$ admits two natural quotients; a quotient $R_\ell^{\mathrm{unr},\psi}$ corresponding to representations which are unramified, and a quotient corresponding to representations for which the ratios of the eigenvalues of Frobenius are equal to ℓ. Because the determinant is fixed, this latter quotient is given by imposing the equation

$$\mathrm{tr}(\rho(\phi))^2 = \ell^{-1}(1+\ell)^2 \psi(\ell).$$

Since $\ell \not\equiv -1 \bmod p$, the right hand side is a unit, and hence there is exactly one square root of this equation which is compatible with the choice of a_ℓ, and so this is equivalent to the equation

$$\mathrm{tr}(\rho(\phi)) = \ell^{-1/2}(1+\ell)\psi^{1/2}(\ell)$$

for the appropriate choice of square root. The ring obtained by imposing this relation on $R_\ell^{\mathrm{univ},\psi,\tau}$ may or may not be either $\mathcal{O}$-flat or reduced, but let $R_\ell^{\mathrm{sp},\psi}$ denote the largest quotient with this property (sp is for special). Its F-points will still include all ramified representations of type τ. The pre-image of the corresponding affine scheme under the projection $\mathrm{Spec}(\widetilde{R}_\ell^{\mathrm{univ}}) \to \mathrm{Spec}(R_\ell^{\mathrm{univ}})$ is equal to $R_\ell^{\mathrm{sp},\psi}[\alpha_\ell]/(\alpha_\ell^2 - \ell^{-1/2}(1+\ell)\psi^{1/2}(\ell)\cdot\alpha_\ell + \psi(\ell))$. The quadratic relation factors as

$$(\alpha_\ell - \psi^{1/2}(\ell)\cdot\ell^{1/2})(\alpha_\ell - \psi^{1/2}(\ell)\cdot\ell^{-1/2}).$$

Define $\widetilde{R}_\ell^{\mathrm{sp},\psi}$ to be the quotient on which $\alpha_\ell = \psi^{1/2}(\ell)\cdot\ell^{-1/2}$. There is a corresponding isomorphism

$$R_\ell^{\mathrm{sp},\psi} \to \widetilde{R}_\ell^{\mathrm{sp},\psi}.$$

On the other hand, the quotient $R_\ell^{\mathrm{unr},\psi}$ is a formally smooth. In this case, we let $\widetilde{R}_\ell^{\mathrm{unr},\psi}$ be the finite flat degree two extension given by adjoining an eigenvalue α_ℓ of the characteristic polynomial of Frobenius. We now let $R_\ell^{\mathrm{mod},\psi} = \widetilde{R}^{\mathrm{univ},\psi,\tau}$ be the image of $\widetilde{R}_\ell^{\mathrm{univ}}$ under the map

$$\widetilde{R}_\ell^{\mathrm{univ}} \to \widetilde{R}_\ell^{\mathrm{unr},\psi} \oplus \widetilde{R}_\ell^{\mathrm{sp},\psi}.$$

It is $\mathcal{O}$-flat and reduced because both $\widetilde{R}_\ell^{\mathrm{unr},\psi}$ and $\widetilde{R}_\ell^{\mathrm{sp},\psi}$ have this property. Moreover, the F-points for finite extensions F/E correspond exactly to either an unramified representation together with a choice of Frobenius, or a ramified ordinary representation together with α_ℓ being sent to the action of Frobenius on the unramified quotient. Geometrically, R_ℓ^{mod} consists of the union of two components, one the special component of $R_\ell^{\mathrm{univ},\psi,\tau}$, and the other a double cover of the unramified component of $R_\ell^{\mathrm{univ},\psi,\tau}$. We also remark that, by construction, the image of the universal deformation ring $R^{\mathrm{univ},\psi}$ in R_ℓ^{mod} will be precisely $R^{\mathrm{univ},\psi,\tau}$.   □

We also note the following:

Corollary 2.3. *Suppose that $p > 2$ and $\ell \neq p$. Let $x : \widetilde{R}_\ell^{\mathrm{univ}} \to E[\epsilon]/\epsilon^2$ be a surjective map so that the image of R^{univ} is E, and the corresponding Galois representation*

$$\rho_x : G_{\mathbf{Q}_\ell} \to \mathrm{GL}_2(E)$$

is unramified. Then x factors through $R_\ell^{\mathrm{mod},\psi}$.

Proof. Such a representation exists exactly when the eigenvalues of $\rho_x(\phi)$ are equal. We see that x certainly factors through $\widetilde{R}^{\mathrm{unr},\psi}$, which is a quotient of $R_\ell^{\mathrm{mod},\psi}$.   □

If the determinant is explicit from the context, we write R_ℓ^{mod} rather than $R_\ell^{\mathrm{mod},\psi}$. We give a precise description of the special fibre of R_ℓ^{mod} when $\ell \equiv 1 \bmod p$, $\psi = 1$, and $\overline{\rho}$ is trivial (this result will only be used for the proof of Theorem 1.3 in Section 4.2.) Note that, since $\overline{\rho}$ by assumption is tamely ramified at ℓ, the image of any deformation also factors through tame inertia, hence through the group $\langle \tau, \phi \rangle$ with $\phi\tau\phi^{-1} = \tau^\ell$.

Lemma 2.4. *Suppose that $p > 2$, that $\ell \equiv 1 \bmod p$, that $\psi|G_{\mathbf{Q}_\ell} \to \mathrm{GL}_2(k)$ is trivial. Then R_ℓ^{mod} represents the functor of deformations of $\overline{\rho}$ to A together with an $\alpha_\ell \in A$ satisfying the following conditions:*

- $\mathrm{Tr}(\rho(\tau)) = 2$,
- $(\rho(\tau) - 1)^2 = 0$,
- $(\rho(\tau) - 1)(\rho(\phi) - \alpha_\ell) = 0$,
- $(\rho(\phi) - \alpha_\ell^{-1})(\rho(\tau) - 1) = 0$,
- $(\rho(\phi) - \alpha_\ell)(\rho(\phi) - \alpha_\ell^{-1}) = 0$.

The argument is similar (but easier) to the corresponding arguments of Snowden ([23, Section 4.5]). In fact, our argument amounts to the case $d = 0$ of a theorem proved by Snowden for all integers $d > 0$. The only reason that Snowden does not consider this case is that, in his context, d is the degree of a finite extension of $\mathbf{Q}_p$.

Proof. The last equation says that α_ℓ satisfies the characteristic polynomial of Frobenius. Hence the functor is certainly represented by a quotient $R_\ell^\dagger$ of the universal such ring $\widetilde{R}_\ell^{\mathrm{univ}}$. Let us show that

$$\mathrm{MaxSpec}(R_\ell^\dagger) = \mathrm{MaxSpec}(R_\ell^{\mathrm{mod}})$$

(inside $\mathrm{MaxSpec}(\widetilde{R}_\ell^{\mathrm{univ}})$). Let $x : R_\ell^\dagger \to F$ be a point of $\mathrm{MaxSpec}(R_\ell^\dagger)$. If $\rho(\tau)$ is trivial, then the equations reduce to the statement that α_ℓ is an eigenvalue of Frobenius, and these correspond exactly to the unramified points of R_ℓ^{mod}. If $\rho(\tau)$ is non-trivial, then, from the first equation, its minimal polynomial will be $(X - 1)^2$, and so, after conjugation, has the shape

$$\rho(\tau) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

The other equations then imply that

$$\rho(\phi) = \begin{pmatrix} \alpha_\ell^{-1} & * \\ 0 & \alpha_\ell \end{pmatrix}.$$

Finally, from the equation $\phi\tau\phi^{-1} = \tau^\ell$, we deduce that $\alpha_\ell^{-2} = \ell$. In particular, the representation is, up to twist, an extension of F by $F(1)$, which corresponds exactly to points on the special component of R_ℓ^{mod}. It follows from [16, Corollary 2.3] (see also Lemma 2.6 below) that $\mathcal{O}$-flat reduced quotients of R_ℓ^{univ} are characterized by their F points for finite extensions F/E. Since we have shown that $R_\ell^\dagger$ and R_ℓ^{mod} have the same such quotients, and since R_ℓ^{mod} is $\mathcal{O}$-flat and reduced, it suffices to show that $R_\ell^\dagger$ is $\mathcal{O}$-flat and reduced. The special

fibre $R_\ell^\dagger/\varpi$ is exactly the completion of $\mathcal{C}_0$ at $c = (1; 1; 0)$ in the notation of [23, Section 3.5]. The proof of this is identical to the proof of [23, Lemma 4.7.4]. On the special fibre, the equation $(\rho(\tau) - 1)^2 = 0$ implies that $(\rho(\tau) - 1)^p = 0$ and so $\rho(\tau^p)$ is trivial, and $\rho(\tau^\ell) = \rho(\tau)$. Hence the action of conjugation by $\rho(\phi)$ on $\rho(\tau)$ is trivial. In [23], the image of inertia factors through an exponent p commutative group which, as a module for $\mathbf{F}_p[\![T]\!]$ where $1 + T$ acts as conjugation by σ, is isomorphic to $U = \mathbf{F}_p \oplus \mathbf{F}_p[\![T]\!]^{\oplus d}$. In our context, the action of inertia commutes with σ and factors through a group $U = \mathbf{F}_p$. In particular, letting $m = \rho(\tau) - 1$ and φ be the image of $\rho(\phi)$, the tuple (φ, α, m) is the corresponding point on $\mathcal{C}_0$. The rest of the argument follows the proof of [23, Theorem 4.7.1]. The ring $\mathcal{C}_0$ has two minimal primes (corresponding to $(\alpha - 1)$ and m, which come from the components $\mathcal{A}_2$ and $\mathcal{B}_0$ respectively, in the notation of [23]). On the other hand, as we have shown, $R_\ell^\dagger[1/\varpi]_{\mathrm{red}} = R^{\mathrm{mod}}[1/\varpi]$ has two minimal primes corresponding to the unramified and ordinary locus, so $R_\ell^\dagger[1/\varpi]$ has two minimal primes, and so, by [23, Propositions 2.2.1 and 2.3.1], it follows that $R_\ell^\dagger$ is $\mathcal{O}$-flat and reduced, and we are done. $\qquad\square$

2.2. The functors D_Q and D_Q^1. In this subsection, $\overline{\rho}$ will be a global Galois representation unramified at p with the primes $N = S \cup P$ as in the introduction. We now define modified deformation rings R_Q^1 and R_Q for certain sets Q of auxiliary primes distinct from N and p. Let D_ℓ denote the decomposition group $G_{\mathbf{Q}_\ell} \subset G_{\mathbf{Q}}$. The superscript 1 refers to weight one, and the lack of superscript will refer to weight p. Note that $R_\emptyset^1 = R^1$. Besides the representation $\overline{\rho}$, part of the data required to define D_Q^1 and D_Q consists of a fixed choice of elements $a_\ell \in k$ for ℓ dividing N and Q. Moreover, for D_Q, we also fix an $a_p \in k$. Let $D_Q^1(A)$ and $D_Q(A)$ consist of deformations ρ to A and a collection of elements $\alpha_\ell \in A$ for $\ell \in N$ (and $\alpha_p \in A$ for $D_Q(A)$) such that:

(i) $\det(\rho) = \psi$, where ψ is the Teichmüller lift of $\det(\overline{\rho})$ for $\rho \in D_Q^1(A)$, $\det(\rho) = \psi \epsilon^{p-1}$ for $\rho \in D_Q(A)$.

(ii) ρ is unramified outside $N \cup Q = S \cup P \cup Q$ for $\rho \in D_Q^1(A)$, and unramified outside $N \cup S \cup Q \cup \{p\}$ for $\rho \in D_Q(A)$.

(iii) If $\ell \in P$, then $\rho|D_\ell \simeq \chi^{-1}\psi|_{D_\ell} \oplus \chi$ if $\rho \in D_Q^1(A)$ and $\rho|D_\ell \simeq \chi^{-1}\psi\epsilon^{p-1}|_{D_\ell} \oplus \chi$ if $\rho \in D_Q(A)$, where χ is an unramified character and $\chi(\mathrm{Frob}_\ell) = \alpha_\ell \equiv a_\ell \bmod \mathfrak{m}$.

(iv) If $\ell \in S$, then $\rho|D_\ell$ corresponds to an A-valued quotient of R_ℓ^{mod}, where we take the determinant to be ψ if $\rho \in D_Q^1(A)$ and $\psi \cdot \epsilon^{p-1}$ if $\rho \in D_Q(A)$, and, in either case $\alpha_\ell \in R_\ell^{\mathrm{mod}}$ is $a_\ell \bmod \mathfrak{m}$.

(v) If $\ell \in Q$, then $\ell \equiv 1 \bmod p$, and $\overline{\rho}(\mathrm{Frob}_\ell)$ has distinct eigenvalues. Then

$$\rho|D_\ell \simeq \phi^{-1}\psi|_{D_\ell} \oplus \phi,$$

where ϕ is a character of $\mathbf{Q}_\ell^\times \subset G_{\mathbf{Q}_\ell}^{\mathrm{ab}}$ such that $\phi(\ell) = \alpha_\ell \equiv a_\ell \bmod \mathfrak{m}$.

(vi) If $\rho \in D_Q(A)$ and $\ell = p$, then $\rho|D_p$ is ordinary with eigenvalue $\alpha_p \equiv a_p \bmod \mathfrak{m}$.

In order for these functors to be non-zero, the a_ℓ for $\ell \in N \cup Q$ must be chosen to be one of the eigenvalues of $\overline{\rho}(\mathrm{Frob}_\ell)$, and a_p must be one of the eigenvalues of $\overline{\rho}(\mathrm{Frob}_p)$. As always, we may extend scalars from k to a field which contains all necessary eigenvalues. For each $\ell \in N$, there exists a corresponding universal framed local deformation ring associated to our deformation problem. There is no subtlety in defining these rings outside the case of primes in S, and at the prime p. The first case was addressed in the previous section. For $\ell = p$, we use

the modified deformation rings as constructed by Snowden ([23], see in particular Section 4.6). For each ℓ, we denote the corresponding modified local deformation ring (with the appropriate determinant) by R_ℓ^{mod}.

Proposition 2.5. *For all of the ℓ different from p, the corresponding modified local deformation ring R_ℓ^{mod} is an $\mathcal{O}$-flat reduced equidimensional ring of relative dimension 3 over $\mathcal{O}$. If $\ell = p$ and $D = D_Q$, then R_ℓ^{mod} is an $\mathcal{O}$-flat reduced equidimensional ring of dimension 4 over $\mathcal{O}$.*

Proof. We consider each deformation ring in turn.

(i) Suppose that $\ell \in P$. By assumption, ψ is ramified at ℓ and hence α_ℓ is uniquely determined by $\rho|D_\ell$. Hence we recover the framed local deformation ring, and the result follows from [5, Lemma 4.11].

(ii) Suppose that $\ell \in S$. Then the result follows from Lemma 2.2.

(iii) Suppose that $\ell \in Q$. The assumption that $\ell \equiv 1 \bmod p$ and that $\bar\rho(\mathrm{Frob}_\ell)$ has no distinct eigenvalues implies that there is no distinction between R_ℓ^{univ} and $\widetilde{R}_\ell^{\mathrm{univ}}$. Moreover, all deformations of $\bar\rho$ will be tamely ramified and split as a direct sum of two characters, and so $R_\ell^{\mathrm{mod}} = R_\ell^{\mathrm{univ},\psi}$ in this case. The ring $R_\ell^{\mathrm{univ},\psi}$ has the desired properties by direct computation, see for example [21, Proposition 7]: it may be identified with $\mathcal{O}[\![X, Y, Z, P]\!]/((1+P)^m - 1)$, where m is the largest power of ℓ dividing $p - 1$.

(iv) If $\ell = p$, and $\bar\rho(\mathrm{Frob}_p)$ has distinct eigenvalues, then the usual definition of an ordinary deformation ring R_p requires a choice of eigenvalue of the unramified quotient, and hence R_p^{mod} is just the usual Kisin ring R_p in this case. If $\bar\rho(\mathrm{Frob}_p)$ has the same eigenvalues, then the local modified deformation ring is exactly the completion of $\mathcal{B}_1$ at $b = (1; 1; 0)$ considered in [23, Section 3.4] and denoted by $\widetilde{R}^\dagger$ in [5, Section 3.7]. The case when $\bar\rho(\mathrm{Frob}_p)$ has the same eigenvalues but is non-scalar corresponds to the localization of $\mathcal{B}_1$ at $((\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}); 1; 0)$. In either case, R_p^{mod} is $\mathcal{O}$-flat, reduced, equidimensional of relative dimension 4 (over $\mathcal{O}$), and Cohen–Macaulay. $\square$

We also present here the following proposition which will be useful later. (Compare [17, Lemma 3.4.12].)

Lemma 2.6. *Let A and B be complete local Noetherian reduced $\mathcal{O}$-flat algebras with residue field k. Then $A \mathbin{\widehat{\otimes}}_{\mathcal{O}} B$ is reduced and $\mathcal{O}$-flat.*

Proof. The $\mathcal{O}$-flatness follows from [12, Section 0, Lemma 19.7.1.2]. Because B is reduced, it follows from [16, Corollary 2.3] that the intersection of the kernels of all morphisms $B \to \mathcal{O}'$ for the ring of integers finite extensions E'/E is trivial. Using this, we may write B as an inverse limit $B = \varprojlim B_i$, where each B_i is reduced and finite flat over $\mathcal{O}$. Then $C = \varprojlim A \otimes B_i$ (now we can replace $\widehat{\otimes}$ by $\otimes$) and it suffices to prove the claim for the usual tensor product when B is finite flat over $\mathcal{O}$, which we now assume. Since C is $\mathcal{O}$-flat, it suffices to show that $C[1/\varpi] = A[1/\varpi] \otimes_E B[1/\varpi]$ is reduced. However, this follows from [1, Chapter V, Section 15.5, Theorem 3 (d)]. $\square$

2.3. Modular curves. Let $N \geq 5$, and let $X_H(N) = X(\Gamma_H(N))$ denote the quotient of $X_1(N)$ by the Sylow p-subgroup of $(\mathbf{Z}/N\mathbf{Z})^\times$ considered as a smooth proper scheme

over $\mathrm{Spec}(\mathcal{O})$ (see [8]). To be precise, the curve $X_H(N)$ is a fine moduli space providing that either $p \geq 5$ or $p = 3$ and N divisible by a prime $q \geq 5$ such that $q \equiv -1 \bmod 3$. This follows either from the computation of stabilizers at the CM points (as in [19, Section 2, p. 64]), or, in the second case, because $X_H(N)$ is a cover of $X_1(q)$. If $p = 3$ and $X_H(N)$ is not a fine moduli space, we simply add a prime $q \geq 5$ and $q \equiv -1 \bmod 3$ to S such that $\bar{\rho}$ is unramified at q. If Q is a collection of auxiliary primes disjoint from N, let $X_H(NQ)$ denote a quotient of $X_1(N)$ by the p-Sylow subgroup of $(\mathbf{Z}/N\mathbf{Z})^\times$ and some subgroup of $(\mathbf{Z}/Q\mathbf{Z})^\times$. (In practice, the cokernel of the corresponding subgroup of $(\mathbf{Z}/Q\mathbf{Z})^\times$ will be a p-group.)

2.4. Hecke algebras. Let ω be the usual pushforward $\pi_* \omega_{\mathcal{E}/X_H(N)}$ of the relative dualizing sheaf along the universal generalized elliptic curve. If A is an $\mathcal{O}$-module, then let $\omega_A^n = \omega^n \otimes_{\mathcal{O}} A$. The (Katz) space of modular forms of weight k and level N is defined to be $H^0(X_H(N), \omega_A^k)$.

We shall now consider a number of Hecke algebras, and discuss the relationship between them. Our coefficient ring or module will either be $A = \mathcal{O}$, $A = E = \mathcal{O} \otimes \mathbf{Q}$, $A = \mathcal{O}/\varpi = k$, $A = \mathcal{O}/\varpi^n$, or $A = E/\mathcal{O}$ unless otherwise specified.

Definition 2.7. *The Hecke algebra $\mathbf{T}_A$ in weight k is the A sub-algebra of*
$$\mathrm{End}_A(H^0(X_H(N), \omega_A^k))$$
generated by the operators T_n for n prime to p and diamond operators $\langle d \rangle$ for d prime to N.

Note that this definition includes the operators T_ℓ for $\ell | N$. These operators can also be denoted by U_ℓ (which is what we shall do below). We now define a variant of these Hecke algebras where we include the Hecke operator at p.

Definition 2.8. *Let $\widetilde{\mathbf{T}}_A$ denote the ring $\mathbf{T}_A$ together with the operator T_p.*

Note that a maximal ideal $\mathfrak{m}$ of $\mathbf{T}$ need no longer a priori be maximal in $\widetilde{\mathbf{T}}$. That is, $\widetilde{\mathbf{T}}_\mathfrak{m}$ will not always be a local ring. However, it will always be a semi-local ring, that is, a direct sum of finitely many local rings.

2.5. Hecke algebras at auxiliary level Q. Let Q be a finite collection of primes congruent to 1 mod p and distinct from N. Let $\mathbf{T}_{Q,A}$ denote the $\mathcal{O}$-algebra generated by Hecke operators away from p acting at level $X_H(NQ)$ with coefficients in A together with diamond operators $\langle d \rangle$ for d prime to N, and let $\widetilde{\mathbf{T}}_{Q,A}$ denote $\mathbf{T}_{Q,A}$ together with the operator T_p.

Suppose that $\bar{\rho}$ is a representation such that the modified deformation ring R is non-zero – equivalently, that $D(k) \neq 0$. Recall that N is equal to the conductor of ψ times the primes in some auxiliary set S which includes (but may be larger than) the set of primes $\ell \nmid p$ where $\bar{\rho}|I_\ell$ is non-trivial and unipotent. In particular, if ℓ divides S, then ℓ divides N exactly once. By Serre's Conjecture [15], any such $\bar{\rho}$ is modular of level $N(\bar{\rho})|N$ and weight p, so we now specialize to the case of weight p, and let $\mathbf{T}_Q = \mathbf{T}_{Q,\mathcal{O}}$. Let $\widetilde{\mathfrak{m}}$ be a maximal ideal of $\widetilde{\mathbf{T}}_Q$ corresponding to $\bar{\rho}$ (and to a choice of a_ℓ for all ℓ dividing N, Q, and p). Let us also suppose that for every prime ℓ dividing Q, the matrix $\bar{\rho}(\mathrm{Frob}_\ell)$ has distinct eigenvalues (since this is an assumption in part (v) of the definition of D_Q).

Proposition 2.9. *There exists a deformation*

$$\rho_Q : G_{\mathbf{Q}} \to \mathrm{GL}_2(\widetilde{\mathbf{T}}_{Q,\widetilde{\mathfrak{m}}})$$

of $\overline{\rho}$ unramified outside pNQ such that $\rho(\mathrm{Frob}_\ell) = T_\ell$ for ℓ prime to pNQ. Let $\rho'_Q = \rho_Q \otimes \eta$, where $\eta^2 = \psi \cdot \epsilon^{p-1} \cdot \det(\rho_Q)^{-1}$. Then ρ'_Q is a deformation of $\overline{\rho}$ in $D_Q(\widetilde{\mathbf{T}}_{Q,\widetilde{\mathfrak{m}}})$. In particular, there is a corresponding map

$$R_Q \to \widetilde{\mathbf{T}}_{Q,\widetilde{\mathfrak{m}}}$$

sending $\mathrm{tr}(\rho^{\mathrm{univ}}(\mathrm{Frob}_\ell)) \in R_Q$ to $\eta(\ell) \cdot T_\ell$ for ℓ not dividing pNQ, sending α_ℓ to $\eta(\ell) \cdot U_\ell$ for ℓ dividing NQ, and sending $\alpha_p + p^{p-1}\psi(p)\alpha_p^{-1}$ to $\eta(p) \cdot T_p$, or equivalently, α_p to the unit root of

$$X^2 - \eta(p) \cdot T_p X + \psi(p) p^{p-1} = 0,$$

which lies in $\widetilde{\mathbf{T}}_{Q,\widetilde{\mathfrak{m}}}$ by Hensel's Lemma.

This proposition is (mostly) an exercise in Atkin–Lehner–Li theory. Indeed, if one assumes that the action of U_ℓ on forms of level $\ell \| N$ is semi-simple (which conjecturally is always the case), then the space of modular forms under consideration will decompose into a direct sum of eigenforms for all the Hecke operators in $\widetilde{\mathbf{T}}_{Q,\mathfrak{m}}$, and then the claim follows immediately from known local-global compatibility for classical modular forms. (The only local–global compatibility we require is given by [7, Theorem 3.1].) In practice, we have to allow for the possibility that U_ℓ may not act semi-simply, although this is not difficult.

Proof. The space of modular forms of weight p is torsion free, so the Hecke algebra is determined by its action on

$$H^0(X_H(NQ), \omega^p)_{\widetilde{\mathfrak{m}}} \otimes E.$$

It suffices to prove the proposition after further decomposing this space into a direct sum of $\widetilde{\mathbf{T}}_{Q,\widetilde{\mathfrak{m}}}$-modules. Enlarging E if necessary, we may assume that all the eigenvalues of all Hecke operators at level dividing NQ are defined over E. Let $\mathbf{T}_Q^{\mathrm{an}}$ denote the anaemic Hecke algebra consisting of endomorphisms of $H^0(X_H(NQ), \omega^p)$ generated by Hecke operators T_n for n prime to pNQ and diamond operators $\langle d \rangle$ for d prime to NQ. There is a map $\mathbf{T}_Q^{\mathrm{an}} \to \widetilde{\mathbf{T}}_Q$; let $\mathfrak{m} \subset \mathbf{T}_Q^{\mathrm{an}}$ denote the inverse of the maximal ideal $\widetilde{\mathfrak{m}}$ (which is determined by $\overline{\rho}$). Note that $\mathfrak{m}$ may correspond to several $\widetilde{\mathfrak{m}}$ in $\widetilde{\mathbf{T}}_Q$; the possible $\widetilde{\mathfrak{m}}$ are indexed by the possible choices of a_ℓ for ℓ dividing pNQ. In any event, there will always be an inclusion:

$$H^0(X_H(NQ), \omega^p)_{\widetilde{\mathfrak{m}}} \otimes E \subset H^0(X_H(NQ), \omega^p)_{\mathfrak{m}} \otimes E.$$

(This would be an equality if we replaced the left hand side with a direct sum over all $\widetilde{\mathfrak{m}}$ which pull back to $\mathfrak{m}$.) The space $H^0(X_H(NQ), \omega^p) \otimes E$ decomposes under $\mathbf{T}_Q^{\mathrm{an}}$ into eigenspaces indexed by newforms f of level dividing NQ. Associated to a cuspidal newform f is a Galois representation ρ_f. In particular, combining all these Galois representations over f with $\overline{\rho}_f = \overline{\rho}$, we obtain a Galois representation

$$\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{T}_{Q,\mathfrak{m}}^{\mathrm{an}} \otimes E).$$

Because the traces of Frobenius elements lie in $\mathbf{T}_{Q,\mathfrak{m}}^{\mathrm{an}}$, and because $\overline{\rho}$ is absolutely irreducible, we may take the image of this Galois representation to land in $\mathrm{GL}_2(\mathbf{T}_{Q,\mathfrak{m}}^{\mathrm{an}})$ by [22, (2.6)]. To this point, we have simply reconstructed the usual construction of the Galois representation into the (anaemic) Hecke algebra.

Let ρ_Q denote the Galois representation induced by composing this with the image of the algebra $\mathbf{T}^{\mathrm{an}}_{Q,\mathfrak{m}}$ in $\widetilde{\mathbf{T}}_{Q,\widetilde{\mathfrak{m}}}$. This will be the ρ_Q of the proposition. The reason for the twist by η is to match the determinant with the required determinant for the functor D_Q. The main point of this proposition is to show that the extra old forms associated to f (with their concomitant actions of U_ℓ) contain exactly the extra information needed to obtain a modified deformation of $\overline{\rho}$ of type D_Q. The eigenspaces corresponding to f will contribute to the localization at $\mathfrak{m}$ if and only if $\overline{\rho}_f = \overline{\rho}$. In particular, the level of f must be divisible by the Serre conductor of $\overline{\rho}$, and hence the level of f is of the form NQ/D, where D is only divisible by primes dividing either S or Q. Hence the integer D is square-free and prime to NQ/D. Suppose that D has d prime divisors. The form f generates a space of 2^d oldforms of level NQ consisting of $f = f(q)$ together with the forms $f(q^m)$ for $m|D$. By Atkin–Lehner–Li theory ([24, Theorem 9.4]) this exhausts the entire space of oldforms associated to f which appear in $H^0(X_H(NQ), \omega^p)_{\mathfrak{m}} \otimes \mathbf{Q}$. Let us now describe the action of U_ℓ on these spaces for ℓ dividing D. Again by Atkin–Lehner, this is given as the tensor product over $\ell|D$ of a two-dimensional space on which U_ℓ acts by the matrix

$$\begin{pmatrix} \mathrm{tr}(\rho_f(\mathrm{Frob}_\ell)) & \langle \ell \rangle \ell^{p-1} \\ -1 & 0 \end{pmatrix}.$$

Here $\mathrm{tr}(\rho_f(\mathrm{Frob}_\ell))$ may also be identified with the eigenvalue of f under the Hecke operator T_ℓ acting at level NQ/D. The element $\mathrm{tr}(\rho_f(\mathrm{Frob}_\ell))$ will lie in the image of $\mathbf{T}^{\mathrm{an}}_Q$ by the Cebotarev density theorem. Note that the eigenvalues of this matrix are precisely the eigenvalues of $\rho_f(\mathrm{Frob}_\ell)$. There are now two possibilities:

(i) The eigenvalues of $\rho_f(\mathrm{Frob}_\ell)$ are distinct. In this case, the space of oldforms over E decomposes further into eigenspaces under U_ℓ. The eigenvalues of U_ℓ will correspond precisely to Galois representations together with a choice of eigenvalue of $\rho_f(\mathrm{Frob}_\ell)$. Each choice of eigenvalue will contribute to the localization at $\widetilde{\mathfrak{m}}$ if and only if the corresponding eigenvalue is $a_\ell \bmod \varpi$. After the global twist to match determinants, such representations will naturally be algebras over R^{mod}_ℓ, where α_ℓ is sent to $\eta(\ell) \cdot U_\ell$ (this follows by the construction of the rings R^{mod}_ℓ, in particular Lemma 2.2 for unramified primes of type S.

(ii) The eigenvalues of $\rho_f(\mathrm{Frob}_\ell)$ are equal. Call the unique eigenvalue b_ℓ. From the explicit matrix description of the action of U_ℓ above, we see that U_ℓ is not a multiple of the scalar matrix, and so it is not diagonalizable. In particular, in the Hecke algebra, the operator U_ℓ satisfies the relation $(U_\ell - b_\ell)^2 = 0$. However, once again (after twisting), there will be a map from R^{mod}_ℓ sending α_ℓ to $\eta(\ell) \cdot U_\ell$, by Corollary 2.3 (the eigenvalues can only be the same for primes of type S).

We remark that the second case above conjecturally never occurs in weight ≥ 2 (see [6]).

Let us now consider the operators U_ℓ for ℓ not dividing D. In this case, the Galois representation ρ_f is ramified at ℓ, and local–global compatibility of Galois representations implies that, after twisting by η, the Galois representation has an unramified quotient on which Frob_ℓ acts via U_ℓ, and hence we have a natural map from R^{mod}_ℓ to $\widetilde{\mathbf{T}}_{Q,\widetilde{\mathfrak{m}}}$ sending (after twisting) α_ℓ to $\eta(\ell) \cdot U_\ell$. Finally, since (by definition) $a_p \in k$ is a unit (it is an eigenvalue of an invertible matrix), the representation ρ_f is ordinary at p, and action of Frobenius on the Galois representations associated to any form f will admit an unramified at p quotient on which Frob_p acts

as U_p. Hence there will be a natural map from R_p^{mod} to $\widetilde{\mathbf{T}}_{Q,\widetilde{\mathfrak{m}}}$ sending α_p to $\eta(p) \cdot U_p$, which is related to T_p via the equation $U_p^2 - T_p U_p + \langle p \rangle p^{p-1} = 0$.   □

Remark 2.10. The theorem above is true in any weight $k \geq 2$, providing that one modifies the definition of D_Q to take into account the weight, and one still works in the ordinary context (so $a_p \in k^\times$).

2.6. Modularity lifting theorems in weight p.   The main goal of this subsection is to prove the following:

Theorem 2.11.   *There is an isomorphism* $R_Q \to \widetilde{\mathbf{T}}_{Q,\widetilde{\mathfrak{m}}}$.

Before proving this theorem, we remark that the "modularity" theorem one can deduce from this $R = \mathbf{T}$ theorem is already well known. In particular, one knows that

$$(R_Q[1/\varpi])^{\mathrm{red}} = (\widetilde{\mathbf{T}}_{Q,\widetilde{\mathfrak{m}}}[1/\varpi])^{\mathrm{red}}$$

(our Hecke algebras will not be reduced if the action of U_ℓ is not semi-simple). Hence the content of this theorem is to upgrade this known result to an integral statement. In order to see how one might do this, note that the modifications of Taylor–Wiles due to Diamond, Kisin, and others ([9, 18]) proceed by constructing a patched module M_∞ over a patched deformation ring R_∞ and a ring of auxiliary diamond operators S_∞. By hook or by crook, one tries to prove that M_∞ is faithful (or nearly faithful) as an R_∞-module. To recover a classical statement, one takes the quotient of R_∞ and M_∞ by the augmentation ideal $\mathfrak{a}$ of S_∞, and recovers the classical ring R and a module M of classical modular forms on which R acts via the quotient $\mathbf{T}$. An essential difficulty, however, is that even if one knows that M_∞ is faithful as an R_∞-module, this does not imply that $M_\infty/\mathfrak{a}$ is faithful as an $R_\infty/\mathfrak{a}$-module; that is, faithfulness is not preserved under quotients. Hence these methods often only allow one to deduce weaker statements concerning reduced quotients. In Wiles' original arguments, however, the auxiliary modules M_D are free over the corresponding Hecke algebras, and one ultimately deduces that the patched module M_∞ is also free over R_∞, from which one can certainly conclude that $M_\infty/\mathfrak{a}$ is free over $R_\infty/\mathfrak{a}$, and hence that $R = \mathbf{T}$. In our argument, we exploit the fact that, by using all the Hecke operators, the multiplicity one theorem for q-expansions allows us to also show that the auxiliary modules M_D are free over certain Hecke algebras, and hence we are able to deduce (as in Wiles) an integral $R = \mathbf{T}$ theorem.

Remark 2.12.   An apology concerning notation: The notation Q that we have used is meant to suggest a collection of Taylor–Wiles primes. Indeed, the primes denoted by Q will play the role of Taylor–Wiles primes in the modularity proof of Section 4.1. However, in the proof of Theorem 2.11 below, the set of primes Q will be fixed, and there will be an auxiliary choice T_D of Taylor–Wiles primes $x \equiv 1 \bmod p^D$. Explicitly, we are proving an $R = \mathbf{T}$ theorem at level p and a level which already includes a fixed collection of Taylor–Wiles primes Q. Hence we require a second auxiliary choice of Taylor–Wiles primes for which we use the letter T rather than Q.

Proof.   We first define a classical unmodified ("natural") deformation ring $R_Q^\natural$ which records deformations which are of the same type as considered in R_Q, except now the extra

choice of eigenvalues is omitted, as is the choice of eigenvalue at $\ell = p$. There is a natural isomorphism

$$R_Q \simeq R_Q^\natural \otimes_{R^{\text{loc}}} R^{\text{mod}},$$

where

$$R^{\text{loc}} := \widehat{\bigotimes_{\ell | pNQ}} R_\ell$$

denotes the corresponding local deformation rings for R^{loc}, and

$$R^{\text{mod}} := \widehat{\bigotimes_{\ell | pNQ}} R_\ell^{\text{mod}}.$$

We remind the reader that one should think about the R^{loc}-algebra R^{mod} as follows: it is the algebra obtained by including the extra information over R^{loc} coming from a choice of Frobenius eigenvalue, and taking a localization of this ring corresponding to fixing (residually) a choice of such an eigenvalue. In particular, the set of components of the generic fibre of R^{mod} is a subset of the components of the generic fibre of R^{loc}. The ring R^{mod} is also reduced by Lemmas 2.5 and 2.6.

We now patch together coherent cohomology modules and we also simultaneously patch Betti cohomology. Namely, we patch the pairs of modules

$$M_D^C = H^0(X_{H_D}(NQ \cdot T_D), \omega_{E/\mathcal{O}}^{\otimes p})_{\widetilde{\mathfrak{m}}}^\vee,$$
$$M_D^B = H^1(X_{H_D}(NQ \cdot T_D), \text{Sym}^{p-2}((E/\mathcal{O})^2))_{\mathfrak{m}}^\vee.$$

The notation C and B refers to coherent and Betti cohomology, respectively. Here T_D is a collection of Taylor–Wiles primes x (distinct from primes dividing NQ) such that $x \equiv 1 \bmod p^D$, and H_D is the subgroup of $(\mathbf{Z}/NQ \cdot T_D\mathbf{Z})^\times$ generated by the kernel of the map

$$(\mathbf{Z}/T_D\mathbf{Z})^\times \to (\mathbf{Z}/p^D\mathbf{Z})^{\#T_D}$$

together with a fixed subgroup of $(\mathbf{Z}/NQ\mathbf{Z})^\times$. The first module has a faithful action of $\widetilde{\mathbf{T}}_{Q \cdot T_D, \widetilde{\mathfrak{m}}}$ and the second has a faithful action of $\mathbf{T}_{Q \cdot T_D, \mathfrak{m}}$. Moreover, the tensor product $M_D^B \otimes_{R^{\text{loc}}} R^{\text{mod}}$ has a natural action of $\widetilde{\mathbf{T}}_{Q \cdot T_D, \widetilde{\mathfrak{m}}}$. We patch together both of these modules for the following reason. The patched Betti cohomology module is known to be nearly faithful over the patched framed natural Galois deformation rings $R^{\text{loc}}[\![x_1, \ldots, x_{r+d-1}]\!]$ by a theorem of Kisin [18] – this essentially amounts to the fact that we already *have* modularity lifting theorems in this context; the goal is to upgrade these theorems to integral statements. On the other hand, the coherent cohomology will be free over the corresponding modified Hecke rings, which allows for an easier passage from patched objects back to finite level.

By Lemma 3.1, the modules M_D^C are free of rank one over $\widetilde{\mathbf{T}}_{Q \cdot T_D, \widetilde{\mathfrak{m}}}$, and so M_D^C/ϖ^D is free of rank one over $\widetilde{\mathbf{T}}_{Q \cdot T_D, \widetilde{\mathfrak{m}}}/\varpi^D$. On the other hand, M_D^B need not be free. However, the action of the Hecke on M_D^B/ϖ^D certainly factors through $\mathbf{T}_{Q \cdot T_D, \mathfrak{m}}/\varpi^D$, and the action of the full Hecke algebra (with Hecke operators for primes dividing N) on $M_D^B/\varpi^D \otimes_{R^{\text{loc}}} R^{\text{mod}}$ factors through $\widetilde{\mathbf{T}}_{Q \cdot T_D, \widetilde{\mathfrak{m}}}/\varpi^D$.

If we patch together all this data simultaneously (together with framings), we deduce (as in the pre-Diamond argument of Wiles) that the patched framed module $M_\infty^{\square, C}$ is free of rank one over $\widetilde{\mathbf{T}}_\infty^\square$ (this is tautological, because each of the modules that is being patched will be free), and that the action of the patched modified Galois deformation ring on $M_\infty^{\square, B} \otimes_{R^{\text{loc}}} R^{\text{mod}}$ acts through $\widetilde{\mathbf{T}}_\infty^\square$. If $R_Q^\square$ denotes the framed version of R_Q over d primes dividing pNQ, then

the Taylor–Wiles method as modified by Kisin gives presentations:

$$R_Q[\![T_1, \ldots, T_{4d-1}]\!] \simeq R_Q^\square \simeq R^{\mathrm{mod}}[\![x_1, \ldots, x_{r+d-1}]\!]/(f_1, \ldots, f_{r+1}),$$

$$R_Q^\natural[\![T_1, \ldots, T_{4d-1}]\!] \simeq R_Q^{\square,\natural} \simeq R^{\mathrm{loc}}[\![x_1, \ldots, x_{r+d-1}]\!]/(f_1, \ldots, f_{r+1}).$$

From [18] we know that $M_\infty^{\square,B}$ is a nearly faithful $R^{\mathrm{loc}}[\![x_1, \ldots, x_{r+d-1}]\!]$-module. From the freeness of M^C over the Hecke algebra, it follows that $\widetilde{\mathbf{T}}_\infty^\square$ is a quotient of the power series ring $R^{\mathrm{mod}}[\![x_1, \ldots, x_{r+d-1}]\!]$. If $\widetilde{\mathbf{T}}_\infty^\square$ is actually isomorphic to this ring, then by taking the quotient by diamond operators, we arrive at the required isomorphism $R_Q \simeq \widetilde{\mathbf{T}}_{Q,\widetilde{\mathfrak{m}}}$. On the other hand, the components of the generic fibre of $R^{\mathrm{mod}}[\![x_1, \ldots, x_{r+d-1}]\!]$ are a subset of those for R^{loc}, so we deduce (cf. [25, Lemma 2.3]) that $\widetilde{\mathbf{T}}_\infty^\square$ is a nearly faithful $R^{\mathrm{mod}}[\![x_1, \ldots, x_{r+d-1}]\!]$-module. Since R^{mod} is reduced and Noetherian, the power series ring has no nilpotent elements, and hence being nearly faithful over this ring is equivalent to being faithful. Thus $\widetilde{\mathbf{T}}_\infty^\square$ is a faithful module, and hence isomorphic to $R^{\mathrm{mod}}[\![x_1, \ldots, x_{r+d-1}]\!]$, and the proof follows. □

Corollary 2.13. *There is an isomorphism* $R_Q/\varpi \simeq \widetilde{\mathbf{T}}_{Q,\widetilde{\mathfrak{m}}}/\varpi \simeq \widetilde{\mathbf{T}}_{Q,k,\widetilde{\mathfrak{m}}}.$

Proof. The first isomorphism is an immediate consequence of the previous theorem, so it suffices to show that $\widetilde{\mathbf{T}}_{Q,\widetilde{\mathfrak{m}}}/\varpi \simeq \widetilde{\mathbf{T}}_{Q,k,\widetilde{\mathfrak{m}}}$, or equivalently, that $\widetilde{\mathbf{T}}_{Q,\widetilde{\mathfrak{m}}}/\varpi$ acts faithfully on $H^0(X, \omega_k^p)$. For this it suffices to note that

$$H^0(X, \omega_k^p)^\vee = (H^0(X, \omega_{E/\mathcal{O}}^p)[\varpi])^\vee = H^0(X, \omega_{E/\mathcal{O}}^p)^\vee/\varpi$$

and that $H^0(X, \omega_{E/\mathcal{O}}^p)^\vee$ is free of rank one over $\widetilde{\mathbf{T}}_{\mathfrak{m}}$ by Lemma 3.1 below. □

3. Katz modular forms

We now study more closely the action of Hecke operators in characteristic p, especially in weights p and one. In this section, we denote $\mathbf{T}_{Q,k}$ and $\widetilde{\mathbf{T}}_{Q,k}$ by $\mathbf{T}$ and $\widetilde{\mathbf{T}}$. We use freely the q-expansion principle, namely, that a form in $H^0(X, \omega_k^n)$ is determined by its image in $k[\![q]\!]$. Multiplication by the Hasse invariant induces a map

$$A : H^0(X, \omega_k) \to H^0(X, \omega_k^p)$$

which is an injection and is the identity map on q-expansions. It follows that this map is $\mathbf{T}$-equivariant, but it is not in general $\widetilde{\mathbf{T}}$-equivariant. There is another map between these spaces induced by the map $q \mapsto q^p$:

$$V : H^0(X, \omega_k) \to H^0(X, \omega_k^p);$$

this map V is also $\mathbf{T}$-equivariant. (Although the corresponding Hecke algebras $\mathbf{T}$ in weights one and p are not the same rings, the meaning of $\mathbf{T}$-equivariance should be clear.)

Lemma 3.1. *Let $\widetilde{\mathfrak{m}}$ be a maximal ideal in $\widetilde{\mathbf{T}}$ in weight p, and let $\mathfrak{m}$ be the corresponding maximal ideal in $\mathbf{T}$. Then* $\dim H^0(X, \omega_k^p)[\widetilde{\mathfrak{m}}] = 1$ *and*

$$\dim H^0(X, \omega_k^p)[\mathfrak{m}] = 1 + \dim H^0(X, \omega_k)[\mathfrak{m}],$$

where $\dim H^0(X, \omega_k)[\mathfrak{m}] \leq 1.$

Proof. The Hecke operators at all primes determine the q-expansion, which proves the first equality. For the second, let $f \in H^0(X, \omega_k^p)[\widetilde{\mathfrak{m}}]$ denote the unique eigenform with leading coefficient 1. There is a homomorphism

$$H^0(X, \omega_k^p)[\mathfrak{m}] \to k[\![q^p]\!] \cap H^0(X, \omega_k^p)[\mathfrak{m}]$$

given by $g \mapsto g - a_1(g)f$. By the q-expansion principle, the kernel of this map is one-dimensional. By the main theorem of Katz [14], the space $k[\![q^p]\!] \cap H^0(X, \omega_k^p)[\mathfrak{m}]$, which lies in the kernel $\ker(\theta)$ of the theta operator, may be identified with the image of $H^0(X, \omega_k)[\mathfrak{m}]$ under V. This gives the first equality. To prove the inequality, we repeat the same argument in weight one, except now [14] implies that $k[\![q^p]\!] \cap H^0(X, \omega_k) = 0$. □

It follows that $\mathbf{T}^1 = \widetilde{\mathbf{T}}^1$ in weight one because $T_p \in \mathbf{T}^1$ and $\mathfrak{m}\mathbf{T}^1 = \widetilde{\mathfrak{m}}\mathbf{T}^1$.

3.1. Doubled modules. We define the notion of a doubled module with respect to $\mathbf{T}$ and $\widetilde{\mathbf{T}}$.

Definition 3.2. *Let $N \subset H^0(X, \omega_k^p)$ be invariant under the action of $\widetilde{\mathbf{T}}$, let*

$$\widetilde{I} = \mathrm{Ann}_{\widetilde{\mathbf{T}}}(N), \quad I = \mathrm{Ann}_{\mathbf{T}}(N) = \widetilde{I} \cap \mathbf{T}.$$

We say that N is doubled if the action of $\widetilde{\mathbf{T}}$ on N acts faithfully through a quotient $\widetilde{\mathbf{T}}/\widetilde{I}$ such that

$$\mathrm{length}(\widetilde{\mathbf{T}}/\widetilde{I}) = 2 \cdot \mathrm{length}(\mathbf{T}/I).$$

Lemma 3.3. *There exists a maximal doubled sub-module of $H^0(X, \omega_k^p)_{\mathfrak{m}}$.*

Proof. In case $\dim H^0(X, \omega_k^p)[\mathfrak{m}] = 1$, then $H^0(X, \omega_k^p)_{\mathfrak{m}}^{\vee}$ is free of rank one over $\mathbf{T}_{\mathfrak{m}}$ and $\widetilde{\mathbf{T}}_{\mathfrak{m}}$, so $\mathbf{T}_{\mathfrak{m}} \simeq \widetilde{\mathbf{T}}_{\mathfrak{m}}$, and the maximal doubled quotient is trivial. Hence we may assume that $\dim H^0(X, \omega_k^p)[\mathfrak{m}] = 2$. By Nakayama's Lemma applied to $\mathbf{T}_{\mathfrak{m}}$, it follows that $\widetilde{\mathbf{T}}_{\mathfrak{m}}$ has rank at most two over $\mathbf{T}_{\mathfrak{m}}$, or equivalently that T_p satisfies a quadratic relation. If N is doubled, however, then $\widetilde{\mathbf{T}}_{\mathfrak{m}}/\widetilde{I}$ must be free of rank two as a $\mathbf{T}_{\mathfrak{m}}/I$-module. In particular, I must act trivially on $\widetilde{\mathbf{T}}_{\mathfrak{m}}/\mathbf{T}_{\mathfrak{m}}$, so it must contain the annihilator of this module. Let J be the annihilator of $\widetilde{\mathbf{T}}_{\mathfrak{m}}/\mathbf{T}_{\mathfrak{m}}$ as a $\mathbf{T}_{\mathfrak{m}}$-module. This is an ideal of $\mathbf{T}_{\mathfrak{m}}$; we claim that it is actually an ideal of $\widetilde{\mathbf{T}}_{\mathfrak{m}}$. By definition, if $a \in \mathbf{T}_{\mathfrak{m}}$ is any element, then a annihilates $\widetilde{\mathbf{T}}_{\mathfrak{m}}/\mathbf{T}_{\mathfrak{m}}$ if and only if it lies in J. Equivalently, we have $ax \in \mathbf{T}_{\mathfrak{m}}$ for all $x \in \widetilde{\mathbf{T}}_{\mathfrak{m}}$ if and only if $a \in J$. To show that J is an ideal of $\widetilde{\mathbf{T}}_{\mathfrak{m}}$, it suffices to show that $aT_p \in J$. By the previous equivalence, we have $aT_p \in \mathbf{T}_{\mathfrak{m}}$. Moreover, since $ax \in \mathbf{T}_{\mathfrak{m}}$ for every $x \in \widetilde{\mathbf{T}}_{\mathfrak{m}}$, we also have $aT_px \in \mathbf{T}_{\mathfrak{m}}$ for every $x \in \widetilde{\mathbf{T}}_{\mathfrak{m}}$. Thus $aT_p \in J$, and J is an ideal of $\widetilde{\mathbf{T}}_{\mathfrak{m}}$. We then observe that $H^0(X, \omega_k^p)[J]$ is doubled, and is thus the maximal doubled sub-module. □

The ideal J is the analogue in this context of the (global) doubling ideal denoted $\mathscr{J}^{\mathrm{glob}}$ in [5].

Let $M \subset H^0(X_k, \omega^p)_{\mathfrak{m}}$ be a maximally doubled module. Hence $M^{\vee}$ is free of rank two over $\mathbf{T}/J$ and free of rank one over $\widetilde{\mathbf{T}}/J$, where $J = \mathrm{Ann}_{\mathbf{T}}(M)$. The only maximal ideal of $\mathbf{T}/J$ is $\mathfrak{m}$, so $\mathbf{T}/J$ is a finite local ring. Let $\ker(\theta)$ denote the subset of elements annihilated by the θ operator, and let $\ker_M(\theta) = \ker(\theta) \cap M$

Lemma 3.4. *The module* $M/\ker_M(\theta)$ *is a faithful* $\mathbf{T}/J$*-module.*

Proof. We have a surjection

$$(\mathbf{T}/J)^2 \simeq M^\vee \to \ker_M(\theta)^\vee.$$

The module $\ker_M(\theta)$ is isomorphic as a $\mathbf{T}$-module to $H^0(X, \omega_k)_\mathfrak{m}$. Hence, by Lemma 3.1, $\ker_M(\theta)^\vee$ is cyclic as a $\mathbf{T}/J$-module. If K denotes the kernel, it follows that $K/\mathfrak{m} \to (\mathbf{T}/\mathfrak{m})^2$ has non-trivial image. Let $x \in K$ denote an element which maps to a non-zero element in $(\mathbf{T}/\mathfrak{m})$. Then the cyclic module in $(\mathbf{T}/J)^2$ generated by x is a faithful $\mathbf{T}/J$-module, and hence K is also a faithful $\mathbf{T}/J$-module. We then have $M/\ker_M(\theta) = K^\vee$. $\square$

Definition 3.5. *There is a* $\mathbf{T}$*-equivariant pairing* $\mathbf{T}/J \times M \to k$ *defined as follows:*

$$\langle T_n, f \rangle = a_1(T_n f).$$

Lemma 3.6. *The map* $\langle *, * \rangle$ *is a perfect pairing between* $\mathbf{T}/J$ *and* $M/\ker_M(\theta)$.

Proof. If $f = \sum a_n q^n$ and $\theta(f) = 0$, then $a_n = 0$ for all $(n, p) = 1$, so $\langle T_n, f \rangle = 0$ for all $T_n \in \mathbf{T}$, and hence for all $T \in \mathbf{T}$. Conversely, if $\langle T_n, f \rangle = 0$, then $a_n = 0$ for all $(n, p) = 1$ and f lies in the kernel of θ. Now suppose that $\langle T, f \rangle = 0$ for all $f \in M$. Since the map is Hecke equivariant, it follows that

$$\langle T_n, Tf \rangle = \langle T T_n, f \rangle = \langle T_n T, f \rangle = \langle T, T_n f \rangle = 0$$

for all n, and hence Tf is trivial in $M/\ker_M(\theta)$. But $\mathbf{T}/J$ acts faithfully on $M/\ker_M(\theta)$, so $T = 0$. $\square$

Lemma 3.7. *The module* $\ker_M(\theta)$ *is a faithful* $\mathbf{T}/J$*-module, and* $\ker_M(\theta)^\vee$ *is free over* $\mathbf{T}/J$ *of rank one.*

Proof. By definition, M is free of rank two over $\mathbf{T}/J$, and so it has the same length as $(\mathbf{T}/J)^2$. However, by Lemma 3.6, $M/\ker_M(\theta)$ and $\mathbf{T}/J$ have the same dimension over k, and hence the same length. It then follows that $\ker_M(\theta)$ has the same length as $\mathbf{T}/J$. Since $\dim \ker_M(\theta)[\mathfrak{m}] = 1$, the module $\ker_M(\theta)^\vee$ is cyclic of the same length as $\mathbf{T}/J$, and thus free of rank one over $\mathbf{T}/J$. Hence $\ker_M(\theta)$ is also faithful as a $\mathbf{T}/J$-module. $\square$

Since, by Lemma 3.1, $\ker(\theta)_\mathfrak{m}^\vee$ is also free of rank one over $\mathbf{T}^1 = \mathbf{T}_{Q,k}^1$, the Hecke algebra in weight one, and $\ker_M(\theta)^\vee$ is a quotient of this module, we deduce the immediate corollary:

Corollary 3.8. *There is a surjection* $\mathbf{T}^1 \to \mathbf{T}/J$.

Now let us fix $X = X_H(NQ)$, and suppose that $\mathfrak{m}$ and $\widetilde{\mathfrak{m}}$ correspond to our residual Galois representation $\bar{\rho}$ together with a suitable choice of a_ℓ.

Proposition 3.9. *There exists a doubled submodule* $M \subset H^0(X_H(NQ), \omega_k^p)_\mathfrak{m}$ *such that the action of* $\mathbf{T}$ *on* M *acts faithfully through* R_Q^1/ϖ. *In particular, there is a surjection* $\mathbf{T}/J \to R_Q^1/\varpi$.

Proof. Let $\widetilde{R}_Q^1$ denote the modification of R_Q where one also takes into account an eigenvalue α_p of $\rho(\mathrm{Frob}_p)$ (recall that representations associated to D_Q^1 are unramified at p). The ring $\widetilde{R}_Q^1$ is a finite flat degree 2 extension of R_Q^1 given as the quotient of $R_Q^1[\alpha_p]$ by the monic quadratic polynomial α_p corresponding to the characteristic polynomial of Frobenius in the universal representation associated to R_Q^1. Let us distinguish two cases. The first is when the eigenvalues of $\overline{\rho}(\mathrm{Frob}_p)$ are distinct, and the second is when they are the same (in the latter case, $\overline{\rho}(\mathrm{Frob}_p)$ may or may not be scalar). Let Σ denote the set of eigenvalues, so $|\Sigma| = 2$ or 1. If $|\Sigma| = 1$, then $\widetilde{R}_Q^1$ is a local ring, and if $|\Sigma| = 2$, it is a semi-local ring with two maximal ideals; indeed, by Hensel's Lemma the quadratic relation satisfied by α_p splits over R_Q^1, and so there is an isomorphism $\widetilde{R}_Q^1 = R_Q^1 \oplus R_Q^1$ in this case. (This is essentially Lemma 2.1.) In particular, the quadratic polynomial has exactly two roots in $\widetilde{R}_Q^1$. There is a surjection

$$\bigoplus_\Sigma R_Q \to \bigoplus_\Sigma R_Q/\varpi \to \widetilde{R}_Q^1/\varpi.$$

Here the sum is over the rings R_Q corresponding to each choice of eigenvalue $a_p \in \Sigma$. The latter map sends α_ℓ to α_ℓ for all ℓ dividing NQ. If $|\Sigma| = 2$, then each α_p goes to the corresponding eigenvalue of Frobenius. If $|\Sigma| = 1$, then α_p goes to α_p. These maps are well defined because, after reduction modulo ϖ, all the local conditions (including the determinant) in the definition of D^1 and D_Q^1 coincide with the exception of $\ell = p$. For $\ell = p$, the enriched ring $\widetilde{R}_Q^1$ receives a map from R_p^{mod}, because α_p in R_p^{mod} is exactly an eigenvalue of Frobenius. Since $\widetilde{\mathbf{T}}_{Q,k,\widetilde{\mathfrak{m}}} \simeq R_Q/\varpi$ by Corollary 2.13, there is a surjection

$$\bigoplus_\Sigma \widetilde{\mathbf{T}}_{Q,k,\widetilde{\mathfrak{m}}} \to \widetilde{R}_Q^1/\varpi.$$

Since $\bigoplus H^0(X_H(NQ), \omega_k^p)_{\widetilde{\mathfrak{m}}}$ is co-free over $\bigoplus \widetilde{\mathbf{T}}_{Q,k,\widetilde{\mathfrak{m}}}$, there certainly exists a module M such that the action of $\widetilde{\mathbf{T}}$ on M is precisely via this quotient $\widetilde{R}_Q^1/\varpi$. Yet this quotient is also finite flat of degree two over R_Q^1/ϖ, which is precisely the image of $\mathbf{T}$ under this map. Hence the submodule M is doubled, and the corresponding action of $\mathbf{T}$ is via R_Q^1/ϖ. The final claim follows from the fact that J is the ideal corresponding to the largest doubled submodule.   □

Remark 3.10.   One alternative way to write this paper was to define the functors D_Q, etc. *without* making a fixed choice of a_ℓ. This would have amounted to replacing the universal local deformation rings R_Q, etc. by universal semi-local deformation rings, which would be isomorphic to a direct sum over all the local rings in this paper and over all possible choices of a_ℓ. We have decided to work with the version of these rings in which choices have been made, however, as evidenced by the proof of the previous proposition, one still has to deal with semi-local rings in some cases at $\ell = p$, because when $\overline{\rho}(\mathrm{Frob}_p)$ has distinct eigenvalues, the corresponding maximal ideal in weight one is determined by the sum $\alpha + \beta$ of these eigenvalues whereas the local rings in higher weight require a choice of α or β.

4. Passage from weight p to weight one

Let $\mathbf{T}_Q^1 = \mathbf{T}_{Q,E/\mathcal{O}}^1$. Recall that the rings $\mathbf{T}^1$, $\widetilde{\mathbf{T}}^1$, $\mathbf{T}$, and $\widetilde{\mathbf{T}}$ of the previous section were abbreviations for the rings $\mathbf{T}_{Q,k}^1$, $\widetilde{\mathbf{T}}_{Q,k}^1$, $\mathbf{T}_{Q,k}$ and $\widetilde{\mathbf{T}}_{Q,k}$ respectively; we return to this expanded notation now.

Corollary 4.1. *There is an isomorphism $R^1_Q/\varpi \to \mathbf{T}^1_{Q,k,\mathfrak{m}} \simeq \mathbf{T}^1_{Q,\mathfrak{m}}/\varpi$.*

Proof. For the first isomorphism, it suffices to note that there exists a map

$$R^1_Q/\varpi \to \mathbf{T}^1_{Q,k,\mathfrak{m}} \to \mathbf{T}_{Q,k,\mathfrak{m}}/J \to R^1_Q/\varpi$$

whose composite is the identity. The existence of the first map follows from the fact that Galois representations in weight one are unramified at p, which follows from [5, Theorem 3.11] (together with the appropriate local–global compatibility away from p, which follows as in the proof of [5, Theorem 3.11] by reduction to characteristic zero in higher weight, together with the proof of Proposition 2.9). The second map comes from Corollary 3.8. The existence of the third map follows from Proposition 3.9. The identification of $\mathbf{T}^1_{Q,k,\mathfrak{m}}$ with $\mathbf{T}^1_{Q,\mathfrak{m}}/\varpi$ follows from the fact that

$$H^0(X_H(NQ), \omega_k)_{\mathfrak{m}} = H^0(X_H(NQ), \omega_{E/\mathcal{O}})_{\mathfrak{m}}[\varpi]$$

and the fact that $H^0(X_H(NQ), \omega_{E/\mathcal{O}})^\vee$ is free over $\mathbf{T}^1_{Q,\mathfrak{m}}$ by Lemma 3.1.　　□

4.1. Modularity in weight one. We have an isomorphism $R^1_Q/\varpi \simeq \mathbf{T}^1_{Q,\mathfrak{m}}/\varpi$ for all collections of Taylor–Wiles primes $Q = Q_D$, and we apply the machinery of [5], in particular Proposition 2.3 as applied in [5, Section 3.8]. We patch the modules $H^0(X_{H_D}(NQ_D), \omega_{E/\mathcal{O}})^\vee_{\mathfrak{m}}$, where Q is a collection of Taylor–Wiles primes $x \equiv 1 \bmod p^D$, and H is the subgroup of $(\mathbf{Z}/NQ_D\mathbf{Z})^\times$ which is generated by the kernel of $(\mathbf{Z}/Q_D\mathbf{Z})^\times \to (\mathbf{Z}/p^D\mathbf{Z})^\times$ and the p-Sylow of $(\mathbf{Z}/N\mathbf{Z})^\times$. We obtain a module M_∞ which is a module over the framed ring of diamond operators $S^\square_\infty$, and a patched deformation ring $R^{1;\square}_\infty$ which is also an algebra over this ring. In contrast to [5], the ring $R^{1;\square}_\infty$ is a power series ring over a completed tensor product

$$R^{\mathrm{mod}} := \widehat{\bigotimes_{\ell | N}} R^{\mathrm{mod}}_\ell,$$

instead of a power series ring over $\mathcal{O}$. Since the modules $H^0(X_Q, \omega_{E/\mathcal{O}})^\vee_{\mathfrak{m}}$ are free over $\mathbf{T}_{Q,\mathfrak{m}}$, the module M_∞ is cyclic as an $R^{1;\square}_\infty$-module. Hence we know that:

(1) M_∞/ϖ is free of rank one over $R^{1;\square}_\infty/\varpi$, as follows from our mod-p modularity results above, in particular Corollary 4.1.

(2) M_∞ is pure of co-dimension one as an $S^\square_\infty$-module; that is, M_∞ is a torsion $S^\square_\infty$-module, and there exists a presentation

$$0 \to (S^\square_\infty)^n \to (S^\square_\infty)^n \to M_\infty \to 0;$$

this is exactly the output of the construction of [5].

The second result is essentially a formal consequence of the method of [5] rather than anything in this paper. This on its own is enough to show that M_∞ will certainly be supported on *some* components of the generic fibre of $R^{1;\square}_\infty$. However, as soon as S contains primes for which $\bar\rho$ is unramified (that is, as soon as we are at non-minimal level), the ring $R^{1;\square}_\infty[1/\varpi]$ will have multiple components. The usual technique for showing that the support of M_∞ is spread over all components is to produce modular lifts with these properties. In our context this is not possible: there are no weight one forms in characteristic zero which are Steinberg at a finite place q (see the proof of Proposition 4.2 below). Our replacement for producing modular points

in characteristic zero is to work on the special fibre, and to show that M_∞/ϖ is (in some sense) spread out as much as possible over $R_\infty^{1,\square}/\varpi$. And we do this (and this is the main point of everything so far) by working in weight p and then descending back to weight one using the doubling method. In particular, we know that M_∞/ϖ is free of rank one over $R_\infty^{1,\square}/\varpi$. From these two properties, we will now deduce that M_∞ is free of rank one over $R_\infty^{1,\square}$, which will imply all our modularity results.

By Nakayama's Lemma, there is certainly an exact sequence of $S_\infty^{\square}$-modules

$$0 \to K \to R_\infty^{1,\square} \to M_\infty \to 0.$$

It suffices to show that $K = 0$. By Nakayama's Lemma again, it suffices to show that $K/\varpi = 0$. Tensoring with $S_\infty^{\square}/\varpi$ (that is, reducing modulo ϖ), we get a long exact sequence

$$\mathrm{Tor}^1(S_\infty^{\square}/\varpi, M_\infty) \to K/\varpi \to R_\infty^{1,\square}/\varpi \xrightarrow{\simeq} M_\infty/\varpi \to 0.$$

Here the last map is an isomorphism by property (1) above. Hence, to prove that K/ϖ is trivial, it suffices to show that

$$\mathrm{Tor}^1(S_\infty^{\square}/\varpi, M_\infty) = M_\infty[\varpi]$$

is trivial. If $M_\infty[\varpi]$ is non-trivial, then, from the purity of M_∞, we claim that M_∞/ϖ will have positive rank over $S_\infty^{\square}/\varpi$. To see this, simply tensor the presentation of M_∞ with $S_\infty^{\square}/\varpi$ to obtain the exact sequence

$$0 \to \mathrm{Tor}^1(S_\infty^{\square}/\varpi, M_\infty) \to (S_\infty^{\square}/\varpi)^n \to (S_\infty^{\square}/\varpi)^n \to M_\infty/\varpi \to 0,$$

from which it follows that

$$\mathrm{rank}_{S_\infty^{\square}/\varpi} M_\infty/\varpi = \mathrm{rank}_{S_\infty^{\square}/\varpi} \mathrm{Tor}^1(S_\infty^{\square}/\varpi, M_\infty).$$

If $\mathrm{Tor}^1(S_\infty^{\square}/\varpi, M_\infty)$ is non-zero, then, as it is a submodule of a free module over $S_\infty^{\square}/\varpi$ which has no associated prime, it also has no associated prime as an $S_\infty^{\square}/\varpi$ module. However, a module with no associated prime over a power series ring over k certainly must have positive rank. This implies that (if $M_\infty[\varpi]$ is non-zero) that M_∞/ϖ has positive rank. Yet this contradicts the fact that $R_\infty^{1,\square}/\varpi \simeq M_\infty/\varpi$ does not have positive rank, as $R_\infty^{1,\square}$ is flat over $\mathcal{O}$ (Lemma 2.6) and of smaller dimension than the ring of diamond operators (by one). Hence M_∞ is free of rank one over $R_\infty^{1,\square}$. But now specializing down to finite level, we deduce that $H^0(X, \omega_{E/\mathcal{O}})_{\mathfrak{m}}^\vee$ is free of rank one over R^1, which proves Theorem 1.2.

4.2. Producing torsion classes. Let

$$f = \sum a_n q^n \in S_1(\Gamma_H(N), \eta)$$

be a cuspidal eigenform of some level N and character η. Let

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

denote the corresponding Artin representation. Assume that $\rho|D_\ell$ is reducible for any prime ℓ dividing N.

Proposition 4.2. *Let f be as above. Let $p > 2$ be a prime such that $\overline{\rho}_f$ is absolutely irreducible and p is prime to the level N of f and the order of η. Then there exists a set of primes ℓ of positive density so that, for each such ℓ, the map*

$$H^0(X(\Gamma_H(N) \cap \Gamma_0(\ell))_{\mathbf{Z}_p}, \omega)_{\mathfrak{m}} \otimes \mathbf{F}_p \to H^0(X(\Gamma_H(N) \cap \Gamma_0(\ell))_{\mathbf{F}_p}, \omega)_{\mathfrak{m}}$$

is not surjective.

Remark 4.3. This implies that Katz' base-change theorem ([13, Theorem 1.7.1]) fails as badly as possible in weight one.

Proof. Suppose, to the contrary, that all such forms of this level lift to characteristic zero. There are no forms in characteristic zero which are new at ℓ of level $\Gamma_0(\ell)$, because any such form would have to be (up to unramified twist) Steinberg at ℓ, and no weight one form in characteristic zero can be Steinberg at any place. The easiest way to see this is that the eigenvalue of U_ℓ would have to be a root of unity times $\ell^{-1/2}$, but this is impossible because Hecke eigenvalues of modular forms are algebraic integers. Hence any Galois representation arising from forms of this level must come from level $\Gamma_H(N)$, and so in particular be unramified at ℓ. Thus, by Theorem 1.2, it suffices to show that there is a non-trivial deformation of $\overline{\rho}$ to the dual numbers which is minimal at N, corresponds to a quotient of R_ℓ^{mod} at the new auxiliary prime ℓ, and is unramified everywhere else. The reduced tangent space of the unramified deformation ring is given by the Selmer group $H^1_\emptyset(\mathbf{Q}, \mathrm{ad}^0(\overline{\rho}))$. Denote the dual Selmer group by $H^1_{\emptyset*}(\mathbf{Q}, \mathrm{ad}^0(\overline{\rho})(1))$. Since these groups are both finite, there exists a finite extension $F/\mathbf{Q}$ which contains the fixed field of $\ker(\rho)$ and such that all the classes in $H^1_{\emptyset*}(\mathbf{Q}, \mathrm{ad}^0(\overline{\rho})(1))$ split completely. Let ℓ be a prime which splits completely in $F(\zeta_p)$. Let $H^1_\Sigma(\mathbf{Q}, \mathrm{ad}^0(\overline{\rho}))$ denote the modified Selmer group where classes are allowed to be arbitrarily ramified at ℓ. By construction of ℓ, the dual Selmer group $H^1_{\Sigma*}(\mathbf{Q}, \mathrm{ad}^0(\overline{\rho}))$ consisting of all dual Selmer classes which split completely at ℓ is equal to $H^1_{\emptyset*}(\mathbf{Q}, \mathrm{ad}^0(\overline{\rho}))$, because the localization map factors through the restriction to G_F, and by construction all classes in the latter group are assumed to split completely over $\mathbf{Q}_\ell$. Hence, the Greenberg–Wiles Euler characteristic formula

$$\frac{|H^1_\Sigma(\mathbf{Q}, \mathrm{ad}^0(\overline{\rho}))|}{|H^1_\emptyset(\mathbf{Q}, \mathrm{ad}^0(\overline{\rho}))|} = \frac{|H^1_\Sigma(\mathbf{Q}, \mathrm{ad}^0(\overline{\rho}))|}{|H^1_{\Sigma*}(\mathbf{Q}, \mathrm{ad}^0(\overline{\rho})(1))|} \cdot \frac{|H^1_{\emptyset*}(\mathbf{Q}, \mathrm{ad}^0(\overline{\rho})(1))|}{|H^1_\emptyset(\mathbf{Q}, \mathrm{ad}^0(\overline{\rho}))|}$$

$$= \frac{|H^1(\mathbf{Q}_\ell, \mathrm{ad}^0(\overline{\rho}))|}{|H^1(\mathbf{F}_\ell, \mathrm{ad}^0(\overline{\rho}))|}$$

$$= |H^0(\mathbf{Q}_\ell, \mathrm{ad}^0(\overline{\rho})(1))|$$

$$= |\mathrm{ad}^0(\overline{\rho})(1)|,$$

the final equality coming from the assumption that ℓ splits completely in $F(\zeta_p)$. Note that $\dim \mathrm{ad}^0(\overline{\rho})(1) = 3 > 0$. It follows that for such choices of ℓ, there exists a deformation

$$\rho : G_\mathbf{Q} \to \mathrm{GL}_2(k[\epsilon]/\epsilon^2)$$

which is minimal at all primes away from ℓ and genuinely ramified at ℓ. Moreover, $\overline{\rho}|D_\ell$ is trivial and $\ell \equiv 1 \bmod p$. It suffices to show that the corresponding deformation arises from a quotient of R_ℓ^{mod}, which was described explicitly in this case by Lemma 2.4. Since we are considering fixed determinant deformations, the trace of the image $\rho(g)$ of any element g is $1 + \det(g) = 2$. However, it is apparent the description of R_ℓ^{mod} in Lemma 2.4 that all the

relations apart from $\mathrm{Tr}(\rho(\tau)) = 2$ lie in $\mathfrak{m}^2$, and so are automatically satisfied for any deformation to $k[\epsilon]/\epsilon^2$. (Note that, associated to ρ, there is a corresponding surjection $R_\ell^{\mathrm{mod}} \to k[\epsilon]/\epsilon^2$ for any choice of $\alpha_\ell \in 1 + \epsilon \cdot k[\epsilon]/\epsilon^2$.) □

4.3. Proof of Theorem 1.3. To prove Theorem 1.3, it suffices to apply Proposition 4.2 to suitably chosen f. Note that the class numbers of the fields $\mathbf{Q}(\sqrt{-23})$ and $\mathbf{Q}(\sqrt{-47})$ are 3 and 5 respectively. This gives rise to suitable weight one forms f with image D_3 and D_5 and level $\Gamma_1(23)$ and $\Gamma_1(47)$ respectively (both with quadratic nebentypus). Applying Proposition 4.2, we deduce the existence of mod-p Katz modular forms which fail to lift for all $p \neq 2, 3, 23$ in the first example and $p \neq 2, 5, 47$ in the second. For $p = 2$, the theorem is known by an example of Mestre [10], completing the proof.

Remark 4.4. The computations of [20] (see also [3]) suggest that Theorem 1.3 should also be true if one insists that f is an *eigenform*. By Serre's conjecture, this would follow if for each $p > 5$ there existed an odd Galois representation: $\overline{\rho} : G_\mathbf{Q} \to \mathrm{GL}_2(k)$ unramified at p with image containing $\mathrm{SL}_2(\mathbf{F}_p)$, although proving this appears difficult.

Remark 4.5. If the cuspform f in Proposition 4.2 is exceptional – that is, the projective image of ρ_f is A_4, S_4, or A_5 –, then the resulting torsion class at level $\Gamma_H(N) \cap \Gamma_0(q)$ will not lift to characteristic zero at any higher level. The reason is that any form g with $\overline{\rho}_f = \overline{\rho}_g$ will have to satisfy $\rho_f = \rho_g$ up to a p-power twist, and the resulting Hecke algebra in weight one cannot give rise to the infinitesimal deformations ρ which arise in the proof of Proposition 4.2.

References

[1] *N. Bourbaki* Éléments de mathématique. Chapitres 4 à 7, Masson, Paris 1981.

[2] *K. Buzzard*, Analytic continuation of overconvergent eigenforms, J. Amer. Math. Soc. **16** (2003), no. 1, 29–55, (electronic).

[3] *K. Buzzard*, Computing weight 1 modular forms over $\mathbf{C}$ and $\overline{\mathbf{F}}_p$, preprint 2013, http://arxiv.org/abs/1205.5077.

[4] *K. Buzzard* and *R. Taylor*, Companion forms and weight one forms, Ann. of Math. (2) **149** (1999), no. 3, 905–919.

[5] *F. Calegari* and *D. Geraghty*, Modularity lifting beyond the Taylor–Wiles method, preprint 2012, http://arxiv.org/abs/1207.4224.

[6] *R. F. Coleman* and *B. Edixhoven*, On the semi-simplicity of the U_p-operator on modular forms, Math. Ann. **310** (1998), no. 1, 119–127.

[7] *H. Darmon, F. Diamond* and *R. Taylor*, Fermat's last theorem, in: Elliptic curves, modular forms & Fermat's last theorem (Hong Kong 1993), International Press, Cambridge (1997), 2–140.

[8] *P. Deligne* and *M. Rapoport*, Les schémas de modules de courbes elliptiques, in: Modular functions of one variable. II (Antwerp 1972), Lecture Notes in Math. **349**, Springer, Berlin (1973), 143–316.

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

[10] *B. Edixhoven*, Comparison of integral structures on spaces of modular forms of weight two, and computation of spaces of forms mod 2 of weight one, J. Inst. Math. Jussieu **5** (2006), no. 1, 1–34.

[11] *T. Gee*, Automorphic lifts of prescribed types, Math. Ann. **350** (2011), no. 1, 107–144.

[12] *A. Grothendieck*, Éléments de géométrie algébrique. IV. Étude locale des schémas et des morphismes de schémas. I, Publ. Math. Inst. Hautes Études Sci. **20** (1964), Paper No. 259.

[13] *N. M. Katz*, p-adic properties of modular schemes and modular forms, in: Modular functions of one variable. III (Antwerp, 1972), Lecture Notes in Math. **350**, Springer, Berlin (1973), 69–190.

[14] *N. M. Katz*, A result on modular forms in characteristic p, in: Modular functions of one variable. V (Bonn 1976), Lecture Notes in Math. **601**, Springer, Berlin (1977), 53–61.

[15] *C. Khare* and *J.-P. Wintenberger*, Serre's modularity conjecture. I, Invent. Math. **178** (2009), no. 3, 485–504.

[16] *C. Khare* and *J.-P. Wintenberger*, Serre's modularity conjecture. II, Invent. Math. **178** (2009), no. 3, 505–586.

[17] *M. Kisin*, Moduli of finite flat group schemes, and modularity, Ann. of Math. (2) **170** (2009), no. 3, 1085–1180.

[18] *M. Kisin*, The Fontaine–Mazur conjecture for GL_2, J. Amer. Math. Soc. **22** (2009), no. 3, 641–690.

[19] *B. Mazur*, Modular curves and the Eisenstein ideal, Publ. Math. Inst. Hautes Études Sci. **47** (1977), 33–186.

[20] *G. J. Schaeffer*, Hecke stability and weight 1 modular forms, preprint 2014, `http://arxiv.org/abs/1406.0408`.

[21] *J. Shotton*, Local deformation rings and a Breuil–Mézard conjecture when $l \neq p$, preprint 2013, `http://arxiv.org/abs/1309.1600`.

[22] *B. de Smit* and *H. W. Lenstra, Jr.*, Explicit construction of universal deformation rings, in: Modular forms and Fermat's last theorem (Boston 1995), Springer, New York (1997), 313–326.

[23] *A. Snowden*, Singularities of ordinary deformation rings, preprint 2011, `http://arxiv.org/abs/1111.3654`.

[24] *W. Stein*, Modular forms, a computational approach, Grad. Stud. Math. **79**, American Mathematical Society, Providence 2007.

[25] *R. Taylor*, Automorphy for some l-adic lifts of automorphic mod l Galois representations. II, Publ. Math. Inst. Hautes Études Sci. **108** (2008), 183–239.

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

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

Frank Calegari, Department of Mathematics, University of Chicago,
5734 S. University Avenue, Room 208C, Chicago, Illinois 60637, USA
e-mail: fcale@math.uchicago.edu