Invent. math. 128, 379–391 (1997)

Inventiones mathematicae

The Taylor-Wiles construction and multiplicity one

Fred Diamond

D.P.M.M.S., University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, UK

Oblatum 30-V-1996 & 13-VIII-1996

1 Introduction

Wiles' proof [17] of the modularity of semistable elliptic curves over **Q** relies on a construction of Taylor and Wiles [16] showing that certain Hecke algebras are complete intersections. These Hecke algebras are defined by considering the action of Hecke operators on spaces of modular forms of "minimal level", or equivalently, on homology groups or Jacobians of modular curves. Taylor and Wiles proceed, roughly speaking, by "patching" algebras arising from forms of different levels.

One of the deep results used in their construction was the fact that the homology of the modular curve becomes a free module (of rank two) over the Hecke algebra upon localization at certain maximal ideals. This result, known as a "multiplicity one" result, is a generalization of a theorem of Mazur [11]. Its proof relied on the q-expansion principle of Deligne-Rapoport and Katz, and the comparison of mod ℓ Betti and de Rham cohomologies (see Sect. 2.1 of [17]).

Multiplicity one was thought to be a crucial ingredient of the Taylor-Wiles construction as well as other parts of Wiles' proof. The purpose of this paper is to explain how to alter the arguments of [16] and [17] so that multiplicity one results are a *byproduct* rather than an *ingredient*.

The key conceptual change underlying this improvement is the following: Rather than prove that (after localization) the Hecke algebra can be identified with the universal deformation ring of a mod ℓ Galois representation, we prove that the homology of the modular curve is a free module over this deformation ring.

To carry this out, we change the Taylor-Wiles construction[1] by 1) "patching" the modules as well as the algebras, and 2) applying the Auslander-Buchsbaum

[1] A similar method was found independently by K. Fujiwara [9] in the course of generalizing some of the results of [16] and [17] to the setting of totally real fields.

formula. The commutative algebra is explained in section in Sect. 2.1 and its application is explained in Sect. 3.1.

In view of other improvements on the methods of [16] and [17], due to Faltings ([16], appendix) and Lenstra [10], we can now conclude that multiplicity one results are actually needed by Wiles only when they are already a consequence of our modified Taylor-Wiles construction.[2] This is discussed in Sect. 3.2. There we also explain how a refinement of [10] given in Sect. 2.2 can be used to obtain multiplicity one results at "non-minimal level."

In addition to simplifying some of the arguments of [16], [17] and [6], the method we describe is expected to make them effective in broader contexts. It can also be used to obtain multiplicity results in situations where one cannot appeal directly to techniques from arithmetic algebraic geometry; we give such an application in Sect. 3.3.

Since the freeness of a certain module turns out to be of more interest than the precise rank, we shall usually speak of "freeness" results from now on, rather than "multiplicity one" results.

2 Freeness criteria

The commutative algebra results which enter into [16] and [17] give criteria for a homomorphism of local rings to be an isomorphism between complete intersections. We now explain how to modify these criteria so that beginning with a module over a local ring, we prove that under certain hypotheses the module is free and the ring is a complete intersection.

2.1 Patching modules

We begin with the key commutative algebra result: a freeness criterion inspired by the isomorphism criterion developed by Wiles, Taylor-Wiles and Faltings (see [16], appendix[3]).

Fix a positive integer r and a finite field k. We consider power series rings $A = k[[S_1, \ldots, S_r]]$ and $B = k[[X_1, \ldots, X_r]]$ and write $\mathfrak{n}$ for the maximal ideal of A.

Theorem 2.1 *Suppose that R is a k-algebra and H is a nonzero R-module, finite-dimensional over k. Suppose that for each positive integer n, there exist k-algebra homomorphisms $\varphi_n : A \to B$ and $\psi_n : B \to R$, a B-module H_n and a B-linear homomorphism $\pi_n : H_n \to H$ such that the following hold:*

(a) ψ_n is surjective and $\psi_n\varphi_n(\mathfrak{n}) = 0$;

[2] Multiplicity one was also initially used by Ribet in his proof of Serre's ϵ-conjecture [14], but it was later removed as an ingredient by his work in [15].

[3] The argument given here can be rephrased in language closer to that of the appendix of [16]. As explained in the introduction, the essential difference is that we patch modules and apply more powerful results from commutative algebra.

(b) π_n *induces an isomorphism* $H_n/\mathfrak{n}H_n \to H$*;*
(c) $\mathrm{Ann}_A H_n = \mathfrak{n}^n$ *and* H_n *is free over* $A/\mathfrak{n}^n$

(where we have regarded H *as a* B*-module via* ψ_n *and* H_n *as an* A*-module via* φ_n*). Then* R *is a complete intersection of dimension zero, and* H *is free over* R*.*

Proof. Let $d = \dim_k H$ and choose a basis $x_1, \ldots, x_d$ for H. For each $n \geq 1$ and $i = 1, \ldots, d$, choose an element $x_{i,n}$ of $\pi_n^{-1}(x_i)$. Thus $x_{1,n}, \ldots, x_{d,n}$ form a basis for H_n over $A/\mathfrak{n}^n A$ by Nakayama's lemma. We consider the homomorphism

$$\mu_n : B \to M_d(A/\mathfrak{n}^n A)$$

of (non-commutative) A-algebras obtained by regarding $(A/\mathfrak{n}^n A)^d \cong H_n$ as a B-module. For each $n \geq 1$ and $j = 1, \ldots, r$, we choose a lift $\nu_n(X_j)$ of $\mu_n(X_j)$ in $M_d(A)$. Since $B^r \times R^r \times M_d(A)^r$ is compact, the sequence

$$(\varphi_n(S_1), \ldots, \varphi_n(S_r), \psi_n(X_1), \ldots, \psi_n(X_r), \nu_n(X_1), \ldots, \nu_n(X_r))$$

has a convergent subsequence (where we have given the $\mathfrak{m}$-adic topology to a complete local Noetherian ring with maximal ideal $\mathfrak{m}$). We denote the limits

$$\varphi_\infty(S_1), \ldots, \varphi_\infty(S_r), \psi_\infty(X_1), \ldots, \psi_\infty(X_r), \nu_\infty(X_1), \ldots, \nu_\infty(X_r).$$

One checks that ν_∞ extends uniquely to a k-algebra homomorphism (in particular, the $\nu_\infty(X_i)$ commute), as of course do φ_∞ and ψ_∞. One finds also that the triangles commute in the resulting diagram

$$\begin{array}{ccccc} & & M_d(A) & & \\ & \nearrow & \uparrow & \searrow & \\ A & \to & B & & M_d(k) \\ & \searrow & \downarrow & \nearrow & \\ & & R & & \end{array}$$

(where $A \to M_d(A)$ and $M_d(A) \to M_d(k)$ are the natural maps, $A \to R$ factors through k, and $R \to M_d(k)$ is defined by the action of R on H). We have now constructed a k-algebra homomorphism $\varphi_\infty : A \to B$ and a B-module H_∞ free of rank d over A. Moreover ψ_∞ is surjective, $\psi_\infty \varphi_\infty(\mathfrak{n}) = 0$ and $H_\infty/\mathfrak{n}H_\infty$ is isomorphic to H as a B-module. Now $\varphi_\infty(S_1), \ldots, \varphi_\infty(S_r)$ is an H_∞-regular sequence, so the B-depth of H_∞ is r. By the Auslander-Buchsbaum-Serre theorem ([1], theorem 2.2.7), H_∞ has finite projective dimension over B. So the Auslander-Buchsbaum formula ([1], theorem 1.3.3),

$$\mathrm{depth}_B H_\infty + \mathrm{proj\ dim}_B H_\infty = \mathrm{depth} B,$$

implies that H_∞ is free over B. It follows that H is free over $B/\varphi_\infty(\mathfrak{n})B$. Therefore ψ_∞ induces an isomorphism

$$B/\varphi_\infty(\mathfrak{n})B \to R$$

and R is a complete intersection of dimension zero.

Remark 2.2 One can ask whether there is an explicit function $f(r,d)$ so that "for each positive integer n" can be replaced by "for some $n > f(r,d)$ where $d = \dim_k H$." See [5] for Rubin's improvement of the Taylor-Wiles-Faltings criterion along such lines.

2.2 A numerical criterion

We now give another freeness criterion, this one based on the numerical isomorphism criterion of Wiles generalized by Lenstra. We first recall the Wiles-Lenstra criterion and then explain how to deduce a freeness criterion from it[4].

We assume that $\mathcal{O}$ is a complete discrete valuation ring with uniformizer λ and residue field k. Fix a complete complete local Noetherian $\mathcal{O}$-algebra R with an $\mathcal{O}$-algebra homomorphism $\pi : R \to \mathcal{O}$. Let $\mathfrak{p}$ denote the kernel of π and let I denote $\mathrm{Ann}_R \mathfrak{p}$.

Theorem 2.3 *Suppose that T is a local Noetherian $\mathcal{O}$-algebra which is finitely generated and free as an $\mathcal{O}$-module. Suppose that $\phi : R \to T$ is a surjective homomorphism of $\mathcal{O}$-algebras with kernel contained in $\mathfrak{p}$. Write π_T for the homomorphism $T \to \mathcal{O}$ such that $\pi_T \phi = \pi$, $\mathfrak{p}_T$ for the kernel of π_T and I_T for $\mathrm{Ann}_T \mathfrak{p}_T$. If $\pi_T(I_T)$ is non-zero, then the following are equivalent:*

1. $\pi_T(I_T) \subset \mathrm{Fitt}_{\mathcal{O}}(\mathfrak{p}/\mathfrak{p}^2)$;
2. $\pi_T(I_T) = \mathrm{Fitt}_{\mathcal{O}}(\mathfrak{p}/\mathfrak{p}^2)$;
3. *R is a complete intersection and ϕ is an isomorphism.*

The theorem was proved by Wiles ([17], appendix) if T is Gorenstein and by Lenstra [10] (see also [5]) in the generality stated here.

We now give the numerical freeness criterion.

Theorem 2.4 *Suppose that H is an R-module, finitely generated and free over $\mathcal{O}$, and that $\mathfrak{p}$ is in the support of H. Let Ω denote $H/(H[\mathfrak{p}_T] + H[I_T])$ where $T = R/\mathrm{Ann}_R H$ (where $\phi : R \to T$ is the natural map and the notation is as in theorem 2.3). Let d denote the rank of $H[\mathfrak{p}]$ over $\mathcal{O}$. If Ω has finite length over $\mathcal{O}$, then the following are equivalent:*

(a) $\mathrm{rank}_{\mathcal{O}} H \leq d \cdot \mathrm{rank}_{\mathcal{O}} T$ *and* $\mathrm{length}_{\mathcal{O}} \Omega \geq d \cdot \mathrm{length}_{\mathcal{O}}(\mathfrak{p}/\mathfrak{p}^2)$;
(b) $\mathrm{rank}_{\mathcal{O}} H = d \cdot \mathrm{rank}_{\mathcal{O}} T$ *and* $\Omega \cong (\mathcal{O}/\mathrm{Fitt}_{\mathcal{O}}(\mathfrak{p}/\mathfrak{p}^2))^d$;
(c) R is a complete intersection and H is free (of rank d) over R.

Proof. The implication (b) $\Rightarrow$ (a) is clear; in fact both inequalities are equalities.

The implication (c) $\Rightarrow$ (b) is immediate from theorem 2.3. Indeed hypothesis (c) implies R is a complete intersection finitely generated and free over $\mathcal{O}$. Note

[4] Theorem 2.1 and the numerical isomorphism criterion already suffice as input from commutative algebra to obtain the main results of [17] without multiplicity one as an ingredient. The freeness criterion below is motivated by a question posed to the author by Lenstra and Ribet; the application is given in Sect. 3.2.

that the $\mathcal{O}$-module $R[\mathfrak{p}] = I$ has rank one since R is Gorenstein. Since we also have $R[I] = \mathfrak{p}$, the desired isomorphism follows from $\pi(I) = \mathrm{Fitt}_{\mathcal{O}}(\mathfrak{p}/\mathfrak{p}^2)$.

We now prove the implication (a) $\Rightarrow$ (c). First note that $H[\mathfrak{p}_T] = H[\mathfrak{p}]$ has $\mathcal{O}$-rank $d > 0$. Since $H/H[I_T]$ is free over $\mathcal{O}$ of rank at most d, it follows that Ω can be generated by d elements as an $\mathcal{O}$-module and hence $H/H[I_T]$ has rank exactly d. We also have $H[\mathfrak{p}_T] \cap H[I_T] = 0$ and

$$\mathrm{Fitt}_{\mathcal{O}}(\mathfrak{p}/\mathfrak{p}^2) \supset \mathrm{Ann}_{\mathcal{O}} H/(H[\mathfrak{p}_T] \oplus H[I_T]) \supset \pi_T(I_T).$$

Since $\mathfrak{p}/\mathfrak{p}^2$ has finite length, so does $\pi_T(I_T)$ and we may apply theorem 2.3 to conclude that ϕ is an isomorphism and R is a complete intersection, finitely generated and free over $\mathcal{O}$.

We also see that $\mathrm{rank}_{\mathcal{O}} H \leq d \cdot \mathrm{rank}_{\mathcal{O}} R$ and Ω is isomorphic to $(\mathcal{O}/\pi(I))^d$. Since IH is contained in $H[\mathfrak{p}]$, the $\mathcal{O}$-module Ω is a quotient of $H/(H[I]+IH)$, which can also be generated by d elements and is annihilated by $\pi(I)$. We therefore conclude that $IH = H[\mathfrak{p}]$ has rank d over $\mathcal{O}$ and H/IH is torsion-free.

Write $\bar{R}$ for $R \otimes_{\mathcal{O}} k$, $\mathfrak{m}$ for its maximal ideal and $\bar{H}$ for $H \otimes_{\mathcal{O}} k$. Since $\bar{R}$ is Gorenstein, we know that $\bar{R}[\mathfrak{m}] = I\bar{R}$ is one-dimensional over k. Since H/IH is torsion-free, the k-vector space $I\bar{H}$ has dimension d. Choose elements $x_1, \ldots, x_d$ of $\bar{H}$ so that $I\bar{H} = \oplus Ix_i$, and let V denote the kernel of the map

$$\begin{array}{rccc} \alpha: & \bar{R}^d & \to & \bar{H} \\ & (r_1, \ldots, r_d) & \mapsto & \sum r_i x_i. \end{array}$$

Since the intersection of V with $I\bar{R}^d$ is trivial, we have $V[\mathfrak{m}] = 0$, hence $V = 0$ and α is injective. Comparing dimensions, we conclude that α is an isomorphism, from which it follows that H is free of rank d over R.

3 Applications of the criteria

Recall that the problem addressed by Wiles in [17] is to prove that a certain ring homomorphism

$$\phi_{\mathscr{D}} : R_{\mathscr{D}} \to \mathbf{T}_{\mathscr{D}}$$

is actually an isomorphism (see also [3], [2] and [6]). Here $R_{\mathscr{D}}$ is the universal deformation ring for a mod ℓ Galois representation

$$\bar{\rho}: \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to GL_2(k)$$

arising from a modular form (the $\mathscr{D}$ refers to the type of deformations we allow), and $\mathbf{T}_{\mathscr{D}}$ is a certain Hecke algebra as described in the introduction. (To define $\mathbf{T}_{\mathscr{D}}$, one uses that $\bar{\rho}$ has a deformation of type $\mathscr{D}$ arising from a modular form, and this relies on the results of Ribet [14] and others.) From its construction we know that the map $\phi_{\mathscr{D}}$ is a surjection of local complete Noetherian $\mathcal{O}$-algebras where $\mathcal{O}$ is the ring of integers of a finite extension of $\mathbf{Q}_\ell$. Moreover $\mathbf{T}_{\mathscr{D}}$ is finite and flat over $\mathcal{O}$. We wish to prove not only that $\phi_{\mathscr{D}}$ is an isomorphism, but that $R_{\mathscr{D}}$ and $\mathbf{T}_{\mathscr{D}}$ are complete intersections.

As our aim here is primarily to illustrate how theorem 2.1 is applied, we shall work only in the more restrictive setting of [3] since the notation, definitions and statements there are better suited for our purposes. More precisely, we explain how to apply theorems 2.1 and 2.4 so that freeness results are a byproduct rather than ingredient in the proof of theorem 3.42 of [3]. This theorem of Wiles states that under certain hypotheses

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

is an isomorphism and these rings are complete intersections, where Σ is now a finite set of primes.

Recall that the theorem is first proved in the case $\Sigma = \emptyset$, then for the case of arbitrary Σ. We discuss these cases separately in Sects. 3.1 and 3.2 below. In Sect. 3.3, we show how theorem 2.1 can be applied to obtain new freeness results in the setting of the homology of Shimura curves.

3.1 Minimal level

Let R denote the k-algebra $R_\emptyset/\lambda R_\emptyset$, where λ is a uniformizer of $\mathcal{O}$ and $k = \mathcal{O}/\lambda\mathcal{O}$. We define H using the homology of the modular curve $X = X_0(N)$ where N is the integer denoted $N_\emptyset$ in (4.2.1) of [3], i.e., N is the Artin conductor of $\bar{\rho}$ (times a factor of ℓ in certain cases). We regard $H_1(X, \mathcal{O})$ as a module for the $\mathcal{O}$-algebra $\mathbf{T}$ generated by the Hecke operators T_m for $m \geq 1$.

Corresponding to $\bar{\rho}$ one finds a maximal ideal $\mathfrak{m}$ of $\mathbf{T}$ and a natural homomorphism

$$\mathbf{T}_\emptyset \to \mathbf{T}_\mathfrak{m} \tag{1}$$

which one checks is an isomorphism (see proposition 4.7 of [3] and remark 3.3 below). We define

$$H = H_1(X, \mathcal{O})^-_\mathfrak{m} \otimes_\mathcal{O} k$$

where the superscript $-$ denotes the set of elements on which complex conjugation acts by -1. We regard H as an R-module via $\phi_\emptyset$.

We shall show that theorem 2.1 applies to R and H *without knowing a priori that* $H_1(X, \mathcal{O})_\mathfrak{m}$ *is free over* $\mathbf{T}_\mathfrak{m}$. Since the action of R factors through $\mathbf{T}_\emptyset/\lambda\mathbf{T}_\emptyset$, we obtain:

Theorem 3.1 *With the above notation and hypotheses:*

(a) $H_1(X, \mathcal{O})_\mathfrak{m}$ *is free over* $\mathbf{T}_\emptyset$*;*
(b) $\mathbf{T}_\emptyset$ *is a complete intersection;*
(c) $\phi_\emptyset$ *is an isomorphism.*

Under our running hypotheses from [3], (a) is due to Mazur and Ribet (see Sect. 5 of [14]), (b) was proved by Taylor and Wiles [16] with (a) as an ingredient, and (c) by Wiles [17] with the first two as ingredients. Faltings showed that the Taylor-Wiles construction could be modified to obtain (b) and (c) simultaneously, but the proof still relied on either (a) or the related q-expansion principle. (See the

appendix of [16] and Sect. 4.3 of [3].) Using theorem 2.1, we can now modify the construction to prove all three assertions at once.

According to theorem 2.49 of [3], the Galois cohomology arguments of [17] and [16] yield an integer r and a finite set of primes Q_n for each positive integer n with the very special properties listed below. In particular, setting $R_n = R_{Q_n}/\lambda R_{Q_n}$, we have that

– R_n is topologically generated as a k-algebra by r elements.

We may therefore define a surjective k-algebra homomorphism $\theta_n : B \to R_n$ where $B = k[[X_1, \ldots, X_r]]$. We also know that

– the cardinality of Q_n is r;
– the primes q in Q_n satisfy $q \equiv 1 \bmod \ell^n$.

We let G_n denote the maximal quotient of $\prod_{q \in Q_n} (\mathbf{Z}/q\mathbf{Z})^\times$ of ℓ-power order. The further conditions that for each q in Q_n:

– $\bar{\rho}$ is unramified at q,
– $\bar{\rho}(\mathrm{Frob}_q)$ has distinct eigenvalues,

allow us to endow R_n with the structure of an algebra over the group ring $k[G_n]$ (see [3], Sect. 2.8). The definition ensures that the image of the maximal ideal of $k[G_n]$ in R is trivial. We also choose a surjective k-algebra homomorphism $A \to k[G_n]$; note that the kernel is contained in $\mathfrak{n}^n$ (in fact, in $\mathfrak{n}^{\ell^n}$), where $\mathfrak{n}$ is the maximal ideal of A. We then define $\varphi_n : A \to B$ so that the diagram

$$\begin{array}{ccc} A & \to & B \\ \downarrow & & \downarrow \\ k[G_n] & \to & R_n \end{array}$$

commutes. Defining ψ_n as the composite of θ_n with $R_n \to R$, hypothesis (a) of theorem 2.1 is satisfied.

To define H_n, we again use the homology of a modular curve. This time, we use the curve X_n corresponding to the group

$$\Gamma_n = \Gamma_0(N) \cap \Gamma_1(M_n)$$

where $M_n = p^2 \prod_{q \in Q_n} q$ for a certain auxiliary prime p (see Sect. 4.3 of [3]). By proposition 4.10 of [3], $\mathbf{T}_{Q_n}$ is isomorphic to the localization at a certain maximal ideal $\mathfrak{m}_n$ of the $\mathscr{O}$-algebra of Hecke operators acting on $H_1(X_n, \mathscr{O})$. We set $\mathbf{T}_n = \mathbf{T}_{Q_n}/\lambda\mathbf{T}_{Q_n}$, and regard

$$H_n' = H_1(X_n, k)^-_{\mathfrak{m}_n}$$

as a B-module via $B \xrightarrow{\theta_n} R_n \to \mathbf{T}_n$, the latter map being induced from ϕ_{Q_n}. The action of G_n then factors through an isomorphism of G_n with the ℓ-Sylow subgroup of $(\mathbf{Z}/M_n\mathbf{Z})^\times \cong \Gamma_0(NM_n)/\Gamma_n$; we abuse notation and denote this subgroup by G_n as well. To obtain H_n and π_n satisfying hypotheses (b) and (c) of theorem 2.1, let $H_n = H_n'/\mathfrak{m}^n H_n'$ and apply the following lemma.

Lemma 3.2 *The $k[G_n]$-module H_n' is free, and there is an isomorphism $H_n'/\mathfrak{n}H_n' \to H$ of $\mathbf{T}_n$-modules.*

This is essentially proved in [16] and [17] using methods of de Shalit [4] and Ribet [12]. We sketch a proof here just to emphasize that it does not rely on freeness results.

Step 1. Let Γ_n' the group corresponding to G_n intermediate to Γ_n and $\Gamma_0(NM_n)$, and let X_n' the corresponding modular curve. Consider the natural map $H_1(X_n, k) \to H_1(X_n', k)$ induced by the projection $X_n \to X_n'$. This is compatible with the Hecke operators and so induces a homomorphism

$$\varpi : H_n' \to H_1(X_n', k)^-_{\mathfrak{m}_n}$$

of $\mathbf{T}_n$-modules. Using that $\bar{\rho}$ is irreducible and hence $\mathfrak{m}_n$ is not Eisenstein, we identify ϖ with the homomorphism

$$H_1(\Gamma_n, k)^-_{\tilde{\mathfrak{m}}_n} \to H_1(\Gamma_n', k)^-_{\tilde{\mathfrak{m}}_n}$$

induced by the inclusion $\Gamma_n \subset \Gamma_n'$, where $\tilde{\mathfrak{m}}_n$ is the preimage of $\mathfrak{m}_n$ in the ring $\mathcal{O}[T_m]_{m\geq 1}$. Appealing to Shapiro's lemma as in the proof of proposition 1 of [16], we see that $H_1(\Gamma_n, k)^-$ is free over $k[G_n]$ and that $H_1(\Gamma_n, k)^-_{G_n}$ maps isomorphically to $H_1(\Gamma_n', k)^-$. It follows that H_n' is free over $k[G_n]$ and ϖ induces an isomorphism

$$H_n'/\mathfrak{m}H_n' \to H_1(\Gamma_n', k)^-_{\mathfrak{m}_n}.$$

Since ℓ divides neither the index of Γ_n' in $\Gamma_0(NM_n)$ nor the order of any elliptic points of $\Gamma_0(NM_n)$, the latter $\mathbf{T}_n$-module is isomorphic to

$$H_1(\Gamma_0(NM_n), k)^-_{\mathfrak{m}_n} \cong H_1(X_0(NM_n), k)^-_{\mathfrak{m}_n}.$$

Step 2. Now we explain how to define a $\mathbf{T}_n$-linear isomorphism

$$H_1(X_0(NM_n), k)_{\mathfrak{m}_n} \to H_1(X_0(N), k)_{\mathfrak{m}}. \tag{2}$$

Let D_n denote the set of positive divisors of M_n. For each d in D_n, we let α_d denote the composite

$$H_1(X_0(NM_n), k) \to H_1(X_0(N), k) \to H_1(X_0(N), k)_{\mathfrak{m}},$$

where the first map is induced by $\tau \mapsto d\tau$ and the second is the natural projection. For d in D_n, we also define elements c_d of $\mathbf{T}_\emptyset$ as follows:

- $c_p = -T_p$ and $c_{p^2} = p$;
- $c_q = -\tilde{\beta}_q$ for $q \in Q_n$, where $\tilde{\beta}_q$ is a chosen root of $X^2 - T_q X + q$ (see Sect. 2.8 of [3]);
- $c_{d_1 d_2} = c_{d_1} c_{d_2}$ for d_1, d_2 relatively prime.

One checks that the map

$$\sum_{d\in D_n} d^{-1}c_d\alpha_d \ : \ H_1(X_0(NM_n),k)\to H_1(X_0(N),k)_{\mathfrak{m}} \tag{3}$$

factors through a $\mathbf{T}_n$-linear map as in (2). The surjectivity is proved as in Sect. 4.5 of [3] using the lemma of Ihara, Ribet and Wiles ([3], lemma 4.28). The injectivity follows from the special properties of p and the primes in Q_n; these properties ensure that the dimensions of the two spaces in (2) coincide.

Remark 3.3 The $\mathcal{O}$-algebra $\mathbf{T}_\emptyset$ can be naturally identified with $\mathbf{T}''_{\mathfrak{m}''}$, where $\mathbf{T}''$ is the $\mathcal{O}$-subalgebra of $\mathbf{T}$ generated by the operators T_p for primes p not dividing $N\ell$ and $\mathfrak{m}''=\mathfrak{m}\cap\mathbf{T}''$. The arguments above can be carried out using $\mathbf{T}''$ instead of $\mathbf{T}$ yielding theorem 3.1 with $H_1(X,\mathcal{O})_{\mathfrak{m}}$ replaced by $H_1(X,\mathcal{O})_{\mathfrak{m}''}$. We thus obtain another proof that (1) is an isomorphism by arguing as follows: We have shown that $H_1(X,\mathcal{O})_{\mathfrak{m}}$ is free over $\mathbf{T}_\emptyset$, so (1) is injective mod λ. Moreover the ranks of $\mathbf{T}_\emptyset$ and $\mathbf{T}_{\mathfrak{m}}$ coincide since the projection $H_1(X,\mathcal{O})_{\mathfrak{m}''}\to H_1(X,\mathcal{O})_{\mathfrak{m}}$ is an isomorphism.

3.2 Non-minimal level

Recall that "multiplicity one" also played a role in Wiles' proof that if $\phi_\emptyset$ is an isomorphism and $\mathbf{T}_\emptyset$ is a complete intersection, then the same holds for ϕ_Σ and $\mathbf{T}_\Sigma$. As in [3], we assume that Σ contains only primes not dividing $N_\emptyset$ (and that if ℓ is in Σ, then $\bar\rho$ is ordinary at ℓ). We define the level N_Σ as in Sect. 4.2 of [3]. Then $\mathbf{T}_\Sigma$ can be identified with the localization of the $\mathcal{O}$-algebra of Hecke operators[5] at level N_Σ at a maximal ideal we denote $\mathfrak{m}_\Sigma$ (see [3], proposition 4.7).

Theorem 3.4 *With the above notation and hypotheses:*

(a) $H_1(X_0(N_\Sigma),\mathcal{O})_{\mathfrak{m}_\Sigma}$ *is free over* $\mathbf{T}_\Sigma$*;*
(b) $\mathbf{T}_\Sigma$ *is a complete intersection;*
(c) ϕ_Σ *is an isomorphism.*

Wiles first proves (a) ([17], Sect. 2.1) generalizing results of Mazur and Ribet. The proof of (b) and (c) in [17] then relies on theorems 3.1 and 2.3 together with a comparison of the "congruence ideals" denoted $\eta_\emptyset$ and η_Σ in [3]. Recall that Wiles proved theorem 2.3 under the hypothesis that T is Gorenstein, and the Gorenstein property for $T=\mathbf{T}_\Sigma$ is established using (a) of the above theorem. Theorem 2.3 was then established by Lenstra [10] without the Gorenstein hypothesis, but to carry out the proof of (b) and (c) without "multiplicity one" as an ingredient, one still needs to compare $\eta_\emptyset$ and η_Σ without using (a). (This comparison can be viewed as a generalization and refinement of Ribet's results

[5] This could be defined as the algebra generated by the operators T_n for all $n\geq 1$. As in remark 3.3 we could instead use a subalgebra generated by fewer Hecke operators, but we do need the operators T_p for primes p in Σ.

in [12].) In Sect. 4.4 of [3] (see especially remarks 4.22 and 4.25), it is observed that the only freeness result needed for the calculation is the one already obtained as assertion (a) of theorem 3.1. We review the argument here, but we use theorem 2.4 instead of theorem 2.3 so as to give a new proof of (a) at the same time.

Enlarging $\mathcal{O}$ if necessary, we may assume there is an $\mathcal{O}$-algebra homomorphism

$$\theta : \mathbf{T}_\emptyset \to \mathcal{O}.$$

(Such a homomorphism is necessarily defined by $\theta(T_p) = a_p(f)$ for some newform f of weight 2, level $N_\emptyset$ and trivial character giving rise to $\bar{\rho}$.) We write π_Σ for the composite $R_\Sigma \to R_\emptyset \to \mathbf{T}_\emptyset \to \mathcal{O}$. We shall show that theorem 2.4 2.4 holds for $H_\Sigma = H_1(X_0(N_\Sigma), \mathcal{O})_{\mathfrak{m}_\Sigma}$ viewed as an R_Σ-module via ϕ_Σ. First note that $\mathfrak{p}_\Sigma = \ker \pi_\Sigma$ is in the support of H_Σ and that the equality

$$\mathrm{rank}_{\mathcal{O}} H_\Sigma = d \cdot \mathrm{rank}\mathbf{T}_\Sigma$$

holds where $d = \mathrm{rank}_{\mathcal{O}} H_\Sigma[\mathfrak{p}_\Sigma] = 2$. We must show that

$$\mathrm{length}_{\mathcal{O}} \Omega_\Sigma \geq 2 \cdot \mathrm{length}_{\mathcal{O}} \mathfrak{p}_\Sigma/\mathfrak{p}_\Sigma^2 \tag{4}$$

where $\Omega_\Sigma = H_\Sigma/(H_\Sigma[\mathfrak{p}_{\mathbf{T}_\Sigma}] + H_\Sigma[I_{\mathbf{T}_\Sigma}])$. By theorems 3.1 and 2.4 we have equality in (4) in the case $\Sigma = \emptyset$. The description of $\mathfrak{p}_\Sigma/\mathfrak{p}_\Sigma^2$ in terms of Galois cohomology leads to the inequality

$$\mathrm{length}_{\mathcal{O}} \mathfrak{p}_\Sigma/\mathfrak{p}_\Sigma^2 \leq \mathrm{length}_{\mathcal{O}} \mathfrak{p}_\emptyset/\mathfrak{p}_\emptyset^2 + \sum_{p \in \Sigma} v_\lambda((p-1)(\theta(T_p)^2 - (p+1)^2))$$

([3], proposition 3.35). It therefore suffices to prove

$$\mathrm{length}_{\mathcal{O}} \Omega_\Sigma \geq \mathrm{length}_{\mathcal{O}} \Omega_\emptyset + 2 \cdot \sum_{p \in \Sigma} v_\lambda((p-1)(\theta(T_p)^2 - (p+1)^2)). \tag{5}$$

A construction similar to the one in Step 2 of the proof of lemma 3.2 yields a $\mathbf{T}_\Sigma$-linear homomorphism

$$\beta_\Sigma : H_\Sigma \to H_\emptyset.$$

The key lemma in the proof of (5) is the following result of Ihara, Ribet and Wiles ([3], lemma 4.24).

Lemma 3.5 *β_Σ is surjective.*

To deduce (5), we use the fact that

$$\mathrm{length}_{\mathcal{O}} \Omega_\Sigma = 2 v_\lambda(\langle x_\Sigma, y_\Sigma \rangle_\Sigma) \tag{6}$$

where $\{x_\Sigma, y_\Sigma\}$ is a basis over $\mathcal{O}$ for $H_\Sigma[\mathfrak{p}_\Sigma]$ and $\langle , \rangle_\Sigma$ is a certain alternating, $\mathbf{T}_\Sigma$-bilinear, $\mathcal{O}$-perfect pairing on H_Σ. (See the proof of lemma 4.17 of [3] and the discussion following it.) Lemma 3.5 implies that β_Σ^t has torsion-free cokernel, where β_Σ^t is the transpose of β_Σ with respect to the pairings $\langle , \rangle_\Sigma$ and $\langle , \rangle_\emptyset$. Therefore β_Σ^t induces an isomorphism $H_\emptyset[\mathfrak{p}_\emptyset] \to H_\Sigma[\mathfrak{p}_\Sigma]$, and (5), with equality in fact, follows from (6) and the computation of the composite $\beta_\Sigma \beta_\Sigma^t$.

3.3 Shimura curves

Recall that the freeness results we obtained in Sects. 3.1 and 3.2 can also proved by a method generalizing Mazur's argument in [11]. The idea of Mazur's argument is that the q-expansion principle yields a multiplicity one result for modular forms in characteristic ℓ, and then comparison theorems from arithmetic algebraic geometry transfer enough information to the homology.

Our variant of the Taylor-Wiles construction can be used to obtain results in situations where one cannot appeal directly to the q-expansion principle. In the setting of the homology of Shimura curves, for example, one can recover some of the freeness results of Ribet [13] and Yang [18]. Their methods combine results obtained from Mazur's argument with ingredients of Ribet's proof of the ϵ-conjecture [14]. In this way they also obtain interesting results where the homology fails to be free over the Hecke algebra. Here we give an example of a freeness result which is disjoint from the ones gotten by the methods of [13] and [18].

We suppose that p and q are distinct primes and D is an indefinite quaternion algebra with center $\mathbf{Q}$ and discriminant pq. We choose a maximal order $\mathscr{O}_D$, a uniformizer π in $\mathscr{O}_{D,p} = \mathscr{O}_D \otimes \mathbf{Z}_p$ and an isomorphism $D \otimes \mathbf{R} \cong M_2(\mathbf{R})$. Write $(p+1) = 2^a m$ with m odd. We let Γ denote the subgroup of $\mathscr{O}_D^\times$ consisting of those elements γ with reduced norm 1 and $\gamma^{2^a} \equiv 1 \bmod \pi\mathscr{O}_{D,p}$. We consider the Shimura curve $X = \Gamma\backslash\mathfrak{H}$ (cf. Sect. 4 of [14] and the introduction of [13]).

Suppose that ℓ is an odd prime not dividing $2pq$ and $\mathscr{O}$ is the ring of integers of a finite extension of $\mathbf{Q}_\ell$ with uniformizer λ and residue field k. We have an action of Hecke operators on the curve's homology $H_1(X, \mathscr{O})$, and we let $\mathbf{T}$ denote the $\mathscr{O}$-algebra generated by the endomorphisms T_s for primes s not dividing pq. Suppose that $\mathfrak{m}$ is a maximal ideal of $\mathbf{T}$ containing ℓ and consider the associated Galois representation

$$\bar{\rho}: \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to GL_2(\mathbf{T}/\mathfrak{m}).$$

(This representation is well-defined up to semi-simplification, and can be constructed using the Jacobian of X.)

We assume that $\bar{\rho}$ has conductor p^2q. In that case the restriction of $\bar{\rho}$ to an inertia group at p is absolutely irreducible. Therefore so is the restriction to G_L where L is the quadratic subfield of $\mathbf{Q}(\mu_\ell)$. In particular we may choose an auxiliary prime r as in lemma 3 of [8]. We fix an isomorphism $\mathscr{O}_D \otimes \mathbf{Z}_r \cong M_2(\mathbf{Z}_r)$ and define

$$\Gamma' = \{\, \gamma \in \Gamma \mid \gamma \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \bmod rM_2(\mathbf{Z}_r) \,\}. \tag{7}$$

From now on we work with Γ' instead of Γ. (Recall that Γ' has no elements of finite order.) We let X' denote the Shimura curve $\Gamma'\backslash\mathfrak{H}$ and consider $H_1(X', \mathscr{O}) \cong H_1(\Gamma', \mathscr{O})$. We define $\mathbf{T}'$ as the $\mathscr{O}$-algebra generated by the endomorphisms T_s of $H_1(X', \mathscr{O})$ for primes s not dividing pqr. We define $\mathfrak{m}'$ as the preimage of $\mathfrak{m}$ in $\mathbf{T}'$ under the natural map $\mathbf{T}' \to \mathbf{T}$ defined by restriction of operators.

Theorem 3.6 *The module $H_1(X', \mathscr{O})_{\mathfrak{m}'}$ is free of rank 8 over $\mathbf{T}'_{\mathfrak{m}'}$.*

This is proved using the methods of Sects. 3.1 and 3.2. We briefly indicate what changes are needed in order to adapt the arguments to the setting of Shimura curves.

Let us first suppose that $p+1$ is not divisible by ℓ. In that case, the deformation ring $R_\emptyset$ is defined as in Sect. 2 of [6]. Since the liftings of $\bar{\rho}$ arising from $\mathbf{T}'_{\mathfrak{m}'}$ are deformations of the type used in the definition of $R_\emptyset$, we obtain a natural homomorphism $R_\emptyset \to \mathbf{T}'_{\mathfrak{m}'}$. We wish to prove that $H = H_1(X', k)_{\mathfrak{m}'}$ is free over $R_\emptyset/\lambda R_\emptyset$. One finds sets of primes Q_n as before and defines subgroups Γ'_n of Γ by congruence conditions analogous to (7) at the primes in Q_n. We then define H'_n as a localization of $H_1(\Gamma'_n, k)$. For "Step 1" of the proof of the analogue of lemma 3.2, one can appeal to the Hochschild-Serre spectral sequence as in the proof of lemma 4.6 of [6]. For "Step 2" of the proof, the surjectivity of the analogue of (2) is gotten from theorem 2 of [7] (and the final argument of [8] in the case $\ell = 3$). The freeness assertion is then deduced as in Sect. 3.1, and the rank is computed after tensoring with $\mathbf{Q}$.

In the case that $p+1$ is divisible by ℓ, the proof relies heavily on the methods of [6]. One finds that the above argument carries over with m replaced by its prime-to-ℓ part and $R_\emptyset$ replaced by its quotient denoted $R^\flat_\emptyset$ in [6]. To complete the proof of the theorem one argues as in Sect. 3.2 (see lemma 4.5 and theorem 5.2 of [6]) with the analogue of (5) taking the form

$$\text{length}_{\mathscr{O}} \Omega_\emptyset \geq \text{length}_{\mathscr{O}} \Omega^\flat_\emptyset + 8v_\lambda(p+1).$$

Acknowledgements. The author is grateful to R. Taylor for helpful discussions regarding this work. Thanks are also due to K. Buzzard, H. Darmon, L. Guo, K. Ribet and the referee for their comments on an earlier version of this paper. The author also benefited from conversations with W. McCallum, G. Pappas and A. Wiles.

This research was supported by the United Kingdom's EPSRC (#GR/J4761). Some of the work was done while visiting Princeton University.

References

1. Bruns, W., Herzog, J.: Cohen-Macaulay Rings, Cambridge Studies in Adv. Math. **39**, Cambridge Univ. Press, Cambridge (1986)
2. Cornell, G., Silverman, J., Stevens, G. (eds.). Proceedings of the Conference on Fermat's Last Theorem, Boston University, August 9–18, 1995
3. Darmon, H., Diamond, F., Taylor, R.: Fermat's Last Theorem, in: Current Developments in Mathematics, 1995, International Press, 1–154
4. de Shalit, E.: On certain Galois representations associated to the modular curve $X_1(p)$, Comp. Math. **95**, 69–100 (1995)
5. de Smit, B., Rubin, K., Schoof, R.: Criteria for complete intersections, to appear in [2]
6. Diamond, F.: On deformation rings and Hecke rings, Annals of Math. **144**, 131–160 (1996)
7. Diamond, F., Taylor, R.: Non-optimal levels of mod ℓ modular representations, Inv. Math. **115**, 435–462 (1994)
8. Diamond, F., Taylor, R.: Lifting modular mod ℓ representations, Duke Math. J. **74**, 253–269 (1994)
9. Fujiwara, K.: Deformation rings and Hecke algebras in the totally real case, (preprint)
10. Lenstra, H.W.: Complete intersections and Gorenstein rings, in: Elliptic Curves, Modular Forms and Fermat's Last Theorem, International Press, Cambridge, 1995, pp. 99–109

11. Mazur, B.: Modular curves and the Eisenstein ideal, Publ. Math. IHES **47**, 33–186 (1977)
12. Ribet, K.: Congruence relations between modular forms, Proc. ICM, 1983, 503–514
13. Ribet, K.: Multiplicities of Galois representations in Jacobians of Shimura curves, in Festschrift of I.I. Piatetski-Shapiro (Part II), Israel Math. Conf. Proc. **3**, 221–236 (1990)
14. Ribet, K.: On modular representations of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ arising from modular forms, Inv. Math. **100**, 431–476 (1990)
15. Ribet, K.: Report on mod ℓ representations of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$, in: Motives, Proc. Symp. in Pure Math. **55**(2), 639–676 (1994)
16. Taylor, R., Wiles, A.: Ring theoretic properties of certain Hecke algebras, Annals of Math. **141**, 553–572 (1995)
17. Wiles, A.: Modular elliptic curves and Fermat's Last Theorem, Annals of Math. **141**, 443–551 (1995)
18. Yang, L.: Multiplicities for Galois representations in the higher weight sheaf cohomology associated to Shimura curves, Ph. D. Thesis, The City University of New York, 1996. thesis

This article was processed by the author using the LaTeX style file *pljour1m* from Springer-Verlag.