In these notes we describe how to apply the Taylor–Wiles method when the numerical coincidence no longer holds. The content of the numerical coincidence is nicely summarized in the opening paragraphs of [CHT08]:

> The method of [TW95] does not extend to GL_n as the basic numerical coincidence on which the method depends (see Corollary 2.43 and Theorem 4.49 of [DDT97]) breaks down. For the Taylor–Wiles method to work when considering a representation
>
> $$r : \mathrm{Gal}(\overline{F}/F) \to G(\overline{\mathbb{Q}}_l)$$
>
> one needs
>
> $$[F : \mathbb{Q}](\dim G - \dim B) = \sum_{v|\infty} H^0(\mathrm{Gal}(\overline{F}_v/F_v), \mathrm{ad}^0\, \overline{r}),$$
>
> where B denotes a Borel subgroup of a (not necessarily connected) reductive group G and ad^0 denotes the kernel of the map, $\mathrm{ad} \to \mathrm{ad}_G$, from ad to its G-coinvariants. This is an 'oddness' condition, which can only hold if F is totally real (or $\mathrm{ad}^0 = (0)$) and r satisfies some sort of self-duality. For instance one can expect positive results if $G = \mathrm{GSp}_{2n}$ or $G = \mathrm{GO}_n$, but not if $G = \mathrm{GL}_n$ for $n > 2$.

The method we discuss here makes no such restriction. We focus on the case of GL_2 over a non-totally real number field F, since all the interesting difficulties appear already in this case. (The case where F is totally real is now very well understood; see for example [Gee].)

Nothing here is original to these notes. Our main reference is Calegari–Geraghty, 'Modularity lifting beyond the Taylor–Wiles method' [CG]. See also Hansen, 'Minimal modularity lifting for GL_2 over an arbitrary number field' [Han]. We also use some ideas from [KT] and [NT16].

These notes come with the following health warning: in order to avoid getting distracted, we have elided certain technical details. In some, but not all, cases, we have included a remark *[in square brackets]* to indicate this.

1. Motivation and Galois deformation theory

Let F be a number field, let p be an odd prime unramified in F, and let Σ denote the set of p-adic places of F. We write F_Σ for the maximal extension of F unramified outside Σ, which lives in some fixed algebraic closure of F. We suppose given as well a finite extension $k/\mathbb{F}_p$ and an absolutely irreducible representation $\overline{\rho} : \mathrm{Gal}(F_\Sigma/F) \to \mathrm{GL}_2(k)$ satisfying the following two conditions:

- The determinant is $\det \overline{\rho} = \epsilon^{-1}$, the inverse of the cyclotomic character.
- For each $v \in \Sigma$, the representation $\overline{\rho}|_{G_{F_v}}$ is finite flat (as usual, G_{F_v} denotes a choice of decomposition group at the place v).

We will study the Galois deformation theory of $\overline{\rho}$. To this end, let $W = W(k)$ denote the Witt vectors of k, and let $\mathcal{C}_W$ denote the category of Artinian local W-algebras with residue field k. We consider the functor $\mathrm{Def}_{\overline{\rho}} : \mathcal{C}_W \to \mathrm{Sets}$, whose value on a ring $A \in \mathcal{C}_W$ is given by the set of liftings $\rho : \mathrm{Gal}(F_\Sigma/F) \to \mathrm{GL}_2(A)$ of $\overline{\rho}$, taken up to strict equivalence, and satisfying the following two conditions:

- The determinant is $\det \rho = \epsilon^{-1}$.
- For each $v \in \Sigma$, the representation $\rho|_{G_{F_v}}$ is finite flat.

According to results of Mazur, the functor $\mathrm{Def}_{\overline{\rho}}$ is pro-represented by a complete Noetherian local W-algebra $R_{\overline{\rho}}$ with residue field k, and there is a natural isomorphism between $\mathrm{Def}_{\overline{\rho}}(k[\epsilon])$ (otherwise said, the Zariski tangent space to the ring $R_{\overline{\rho}}$) and the k-vector space

$$H^1_{\mathcal{L}}(F_\Sigma/F, \mathrm{ad}^0\,\overline{\rho}) \subset H^1(F_\Sigma/F, \mathrm{ad}^0\,\overline{\rho}),$$

otherwise called the Selmer group. The subscript $\mathcal{L}$ indicates that we have chosen a set of local conditions $\mathcal{L}_v \subset H^1(F_v, \mathrm{ad}^0\,\overline{\rho})$ for the places $v \in \Sigma$; these reflect the fact that we are working only with those deformations of $\overline{\rho}$ which are finite flat at p.

Writing $g = \dim_k H^1_{\mathcal{L}}(F_\Sigma/F, \mathrm{ad}^0\,\overline{\rho}) = h^1_{\mathcal{L}}(\mathrm{ad}^0\,\overline{\rho})$, say, we can therefore find a surjection $W[\![X_1, \ldots, X_g]\!] \to R_{\overline{\rho}}$ of W-algebras which induces an isomorphism on Zariski tangent spaces.

According to results of Wiles, there is an equation

$$h^1_{\mathcal{L}}(\mathrm{ad}^0\,\overline{\rho}) = h^1_{\mathcal{L}^\perp}(\mathrm{ad}^0\,\overline{\rho}(1)) - r_2,$$

where $r = h^1_{\mathcal{L}^\perp}(\mathrm{ad}^0\,\overline{\rho}(1)) = \dim_k H^1_{\mathcal{L}^\perp}(F_\Sigma/F, \mathrm{ad}^0\,\overline{\rho}(1))$ denotes the dual Selmer group, defined with respect to the dual local conditions $\mathcal{L}_v^\perp \subset H^1(F_v, \mathrm{ad}^0\,\overline{\rho}(1))$, and we write r_1, r_2 for the numbers of real and complex places of F, respectively. *[This formula follows from the Euler characteristic formula of Wiles, together with a calculation of the relevant local terms. It uses that $\overline{\rho}$ is odd at the real places, and we have assumed that any terms which vanish when $\overline{\rho}$ is sufficiently non-degenerate are indeed 0. See [DDT97, Theorem 2.19].]*

The 'numerical coincidence' which occurs in the context of the 'classical' Taylor–Wiles method is the equality $h^1_{\mathcal{L}}(\mathrm{ad}^0\,\overline{\rho}) = h^1_{\mathcal{L}^\perp}(\mathrm{ad}^0\,\overline{\rho}(1))$; with the assumptions in effect here, this just means that $r_2 = 0$, i.e. that the number field F is totally real. We now attempt to explain why this is relevant.

By definition, a Taylor–Wiles set for $\overline{\rho}$ is a set $Q = \{v_1, \ldots, v_r\}$ of finite places of the number field F satisfying the following conditions:

(i) For each $i = 1, \ldots, r$, we have $q_{v_i} \equiv 1 \bmod p$. (By definition, q_v is the size of the residue field $k(v)$ at the place v.)

(ii) For each $i = 1, \ldots, r$, $\overline{\rho}(\mathrm{Frob}_{v_i})$ has 2 distinct eigenvalues $\alpha_i, \beta_i \in k$.

(iii) The natural localization map $H^1_{\mathcal{L}^\perp}(F_\Sigma/F, \mathrm{ad}^0\,\overline{\rho}(1)) \to \oplus_{i=1}^r H^1(F_{v_i}, \mathrm{ad}^0\,\overline{\rho}(1))$ is injective and $r = h^1_{\mathcal{L}^\perp}(\mathrm{ad}^0\,\overline{\rho}(1))$. (The image lies in the unramified subspace, so another way to say this is that the localization map

$$H^1_{\mathcal{L}^\perp}(F_\Sigma/F, \mathrm{ad}^0\,\overline{\rho}(1)) \to \bigoplus_{v \in Q} H^1(k(v), \mathrm{ad}^0\,\overline{\rho}(1))$$

is an isomorphism.)

One can show, using the Chebotarev density theorem, that there are many Taylor–Wiles sets for $\overline{\rho}$. *[In fact, this requires that $\overline{\rho}$ has big image, in some sense. It suffices to assume that $p > 5$ and that $\overline{\rho}|_{G_{F(\zeta_p)}}$ is absolutely irreducible.]* If Q is a fixed choice of Taylor–Wiles set, then we define a new deformation functor $\mathrm{Def}_{\overline{\rho},Q} : \mathcal{C}_W \to \mathrm{Sets}$ by sending a ring $A \in \mathcal{C}_W$ to the set of strict equivalence classes of liftings $\rho : \mathrm{Gal}(F_{\Sigma \cup Q}/F) \to \mathrm{GL}_2(A)$ which have determinant ϵ^{-1}, are finite flat at p, and satisfy no extra condition at the places of Q. Again, one can show that this functor is pro-represented by a complete Noetherian local W-algebra $R_{\overline{\rho},Q}$ with residue field k.

We now describe the key properties of the ring $R_{\overline{\rho},Q}$. The first property is that it has a natural structure of $W[\Delta_Q]$-algebra, where Δ_Q is the maximal p-power order quotient of the group $\prod_{i=1}^r k(v_i)^\times$. (Note that each of the groups $k(v_i)^\times$ is cyclic of order divisible by p, because

$q_{v_i} \equiv 1 \bmod p$. Thus if $q_{v_i} \equiv 1 \bmod p^n$ for each $i = 1, \dots, r$, then there is an isomorphism $\Delta_Q/(p^n) \cong (\mathbb{Z}/p^n\mathbb{Z})^r$.) This structure arises because for each $i = 1, \dots, r$ we can decompose the universal lifting of type $\mathrm{Def}_{\overline{\rho},Q}$ as

$$\rho_Q^{\mathrm{univ}}|_{G_{F_{v_i}}} = A_i \oplus B_i,$$

where $A_i, B_i : G_{F_{v_i}} \to R_{\overline{\rho},Q}^\times$ are continuous characters which are unramified (and send Frobenius to α_i or β_i, respectively) modulo the maximal ideal. By local class field theory, the character A_i determines a homomorphism $k(v_i)^\times \to R_{\overline{\rho},Q}^\times$ of p-power order, and taking the product of these characters determines the desired homomorphism $\Delta_Q \to R_{\overline{\rho},Q}^\times$. It is an easy consequence of the definitions that the natural map $R_{\overline{\rho},Q} \to R_{\overline{\rho}}$ factors through an isomorphism

$$R_{\overline{\rho},Q} \otimes_{W[\Delta_Q]} W \cong R_{\overline{\rho}}.$$

The second property is that the natural surjection $R_{\overline{\rho},Q} \to R_{\overline{\rho}}$ (induced by the inclusion $\mathrm{Def}_{\overline{\rho}} \subset \mathrm{Def}_{\overline{\rho},Q}$) induces an isomorphism of tangent spaces, implying the existence of a surjective homomorphism $W[\![X_1, \dots, X_g]\!] \to R_{\overline{\rho},Q}$ of W-algebras. This fact again follows from Wiles' Euler characteristic formula, and is the reason for including property (iii) in the definition of a Taylor–Wiles set.

We can now explain the basic idea of the Taylor–Wiles method, as it appears e.g. in the original papers of Wiles and Taylor–Wiles. The calculation of the dimension of tangent spaces describes a uniform upper bound for the size of the rings $R_{\overline{\rho},Q}$, as the set Q varies: they are all quotients of $W[\![X_1, \dots, X_g]\!]$. The rings $W[\Delta_Q]$ have Zariski tangent space of dimension r, and can be used to give a lower bound for the size of the part of $R_{\overline{\rho},Q}$ coming from spaces of modular forms if there is e.g. a compatible action of $W[\Delta_Q]$ by diamond operators on a suitable space of modular forms, using that as Q varies, the rings Δ_Q approximate more and more closely the ring $W[\![\mathbb{Z}_p^r]\!] \cong W[\![S_1, \dots, S_r]\!]$. When the 'numerical coincidence' $g = r$ holds, these upper and lower bounds agree and (after a lot more work!) one can show that $R_{\overline{\rho}}$ is exhausted by Galois representations arising from modular forms (in other words, that $R = \mathbb{T}$).

2. Cohomology and automorphic forms

We now introduce some Hecke algebras, continuing with the notation of §1: F is a number field, p is an odd prime unramified in F, and k is a finite field of characteristic p.

Let $G = \mathrm{PGL}_{2,F}$, and fix a choice of maximal compact subgroup $U_\infty \subset G(F \otimes_{\mathbb{Q}} \mathbb{R})$. If $U = \prod_v U_v \subset \mathrm{GL}_2(\mathbb{A}_F^\infty)$ is a compact open subgroup, then we define a topological space as a double quotient

$$X_U = G(F) \backslash G(\mathbb{A}_F) / U U_\infty.$$

We can write down an isomorphism $X_U \cong \coprod_{i=1}^n \Gamma_i \backslash \mathbb{H}_2^{r_1} \times \mathbb{H}_3^{r_2}$, where $\mathbb{H}_d$ denotes hyperbolic space of dimension d and the Γ_i are arithmetic subgroups of $\mathrm{PGL}_2(F \otimes_{\mathbb{Q}} \mathbb{R})$ which depend on U. If the groups Γ_i are neat, then we see that X_U has a natural structure of smooth manifold of dimension $2[F : \mathbb{Q}] - r_2$. *[We will assume that U always satisfies this neatness condition, and that X_U is a smooth manifold, in what follows. This is a minor technical point that does not cause any great difficulty.]* Then the cohomology groups $H^*(X_U, W)$ are finitely generated W-modules, non-zero only in the range $[0, 2[F : \mathbb{Q}] - r_2]$.

It is usual to consider the action of Hecke operators on the groups $H^*(X_U, W)$. We are instead going to construct Hecke algebras in the derived category, following [NT16]. We write

$\mathbf{K}(W)$ for the category of chain complexes of W-modules, and $\mathbf{D}(W)$ for the derived category of chain complexes: it has the same objects as $\mathbf{K}(W)$, but all quasi-isomorphisms have been inverted. The cohomology groups $H^*(X_U, W)$ are computed as the cohomology groups of a complex $C(X_U, W) \in \mathbf{D}(W)$, which is well-defined up to isomorphism.

Hecke operators exist in the derived category: there is a natural map

$$\mathcal{H}(\mathrm{GL}_2(\mathbb{A}_F^\infty), U) \to \mathrm{End}_{\mathbf{D}(W)}(C(X_U, W)),$$

where $\mathcal{H}(\mathrm{GL}_2(\mathbb{A}_F^\infty), U)$ is the usual double coset algebra. If $S \supset \Sigma$ is a finite set of primes of F and $U_v = \mathrm{GL}_2(\mathcal{O}_{F_v})$ for all primes $v \notin S$ of F, then we can define a Hecke algebra $\mathbb{T}_U$ as the W-subalgebra of $\mathrm{End}_{\mathbf{D}(W)}(C(X_U, W))$ generated by the usual unramified Hecke operators $T_v = [\mathrm{GL}_2(\mathcal{O}_{F_v}) \operatorname{diag}(\varpi_v, 1) \mathrm{GL}_2(\mathcal{O}_{F_v})]$ for primes $v \notin S$. These operators commute in $\mathcal{H}(\mathrm{GL}_2(\mathbb{A}_F^\infty), U)$, so we see that $\mathbb{T}_U$ is a finite commutative W-algebra. In particular, it has finitely many maximal ideals $\mathfrak{m}$, each of which has residue field $\mathbb{T}_U/\mathfrak{m}$ a finite extension of k. *[The algebra $\mathbb{T}_U$ depends a priori on the choice of set S. However, a standard argument (see [KT, Lemma 6.20]) shows that if we assume Conjecture A below, and $\mathfrak{m} \subset \mathbb{T}_U$ is a non-Eisenstein maximal ideal, then the localization $\mathbb{T}_{U,\mathfrak{m}}$ is independent of the choice of S. We therefore feel free to ignore this dependence in our notation.]*

Let $\mathbb{T}'_U$ denote the quotient of $\mathbb{T}_U$ which acts faithfully on cohomology. Since the complex $C(X_U, W)$ is filtered by its finitely many cohomology objects, the surjection $\mathbb{T}_U \to \mathbb{T}'_U$ has nilpotent kernel; in particular, every maximal ideal of $\mathbb{T}_U$ appears in the support of $H^*(X_U, W)$. (See e.g. [KT, Lemma 2.5].)

To get started, we are going to state the two basic conjectures that Calegari–Geraghty need to get their generalization of the Taylor–Wiles method to work. The first is about existence of Galois representations with $\mathbb{T}_U$-coefficients; the second is about the behaviour of the cohomology groups $H^*(X_U, W)$.

Conjecture A. (i) *For each maximal ideal $\mathfrak{m} \subset \mathbb{T}_U$, there exists a continuous representation $\overline{\rho}_\mathfrak{m} : \mathrm{Gal}(F_S/F) \to \mathrm{GL}_2(\mathbb{T}_U/\mathfrak{m})$ such that $\det \overline{\rho}_\mathfrak{m} = \epsilon^{-1}$ and for all $v \notin S$, $\operatorname{tr} \overline{\rho}_\mathfrak{m}(\mathrm{Frob}_v) = T_v$ mod $\mathfrak{m}$.*

(ii) *If $\mathfrak{m} \subset \mathbb{T}_U$ is a maximal ideal such that $\overline{\rho}_\mathfrak{m}$ is absolutely irreducible (in which case we say that $\mathfrak{m}$ is non-Eisenstein), there exists a lifting $\rho_\mathfrak{m} : \mathrm{Gal}(F_S/F) \to \mathrm{GL}_2(\mathbb{T}_{U,\mathfrak{m}})$ of $\overline{\rho}_\mathfrak{m}$ such that $\det \rho_\mathfrak{m} = \epsilon^{-1}$ and for all $v \notin S$, $\operatorname{tr} \rho_\mathfrak{m}(\mathrm{Frob}_v) = T_v$.*

(iii) *If $\mathfrak{m}$ is a non-Eisenstein maximal ideal and $U_v = \mathrm{GL}_2(\mathcal{O}_{F_v})$ for all $v \in \Sigma$, then the representation $\rho_\mathfrak{m}$ is finite flat at p.*

Suppose that $U = \prod_v \mathrm{GL}_2(\mathcal{O}_{F_v})$ and $\mathfrak{m} \subset \mathbb{T}_U$ is non-Eisenstein. Then $\overline{\rho} = \overline{\rho}_\mathfrak{m}$ satisfies the conditions of the previous section, and the lifting $\rho_\mathfrak{m}$ determines a surjective homomorphism $R_{\overline{\rho}} \to \mathbb{T}_{U,\mathfrak{m}}$. Our goal in these notes is to prove that this map is in fact an isomorphism.

Conjecture B. *Let $\mathfrak{m} \subset \mathbb{T}_U$ be a non-Eisenstein maximal ideal. Then the groups $H^*(X_U, k)_\mathfrak{m}$ are concentrated in the range $[r_1, r_1 + r_2]$.*

This conjecture is motivated by what happens rationally: the groups $H^*(X_U, W)[1/p]$ can be computed in terms of cuspidal cohomological automorphic forms, and these indeed contribute non-zero classes exactly in degrees in the range $[r_1, r_1 + r_2]$.

We will assume both of these conjectures in what follows.

3. The classical case

Having set the scene, we now describe the Taylor–Wiles argument in the classical case where the base number field F is totally real and the numerical coincidence $g = r$ is in effect. The version of the argument we describe here is in fact a modification of the original arguments of Taylor and Wiles which is due to Diamond. We observe that the case where F is totally real is the only case currently where both Conjectures A and B are known. *[More precisely, Conjecture B is known in many cases thanks to work of Dimitrov [Dim05]; in practice, one would use the Jacquet–Langlands correspondence to move to a form of* PGL_2 *(e.g. definite at infinity) for which Conjecture B becomes trivial.]*

We thus take F to be totally real, $U = \prod_v \mathrm{GL}_2(\mathcal{O}_{F_v})$, and $\mathfrak{m} \subset \mathbb{T}_U$ a non-Eisenstein maximal ideal. Let $\overline{\rho} = \overline{\rho}_{\mathfrak{m}}$. We will show that the map $R_{\overline{\rho}} \to \mathbb{T}_{U,\mathfrak{m}}$ is an isomorphism. In this situation, X_U is the set of complex points of a Hilbert modular variety, and the groups $H^i(X_U, W)_{\mathfrak{m}}$ are free W-modules, non-zero only in the middle degree $i = [F : \mathbb{Q}]$. This implies that the Hecke algebra $\mathbb{T}_{U,\mathfrak{m}}$ acts faithfully on $H_0 = H^{[F:\mathbb{Q}]}(X_U, W)_{\mathfrak{m}}$, so we can forget about derived categories for this section. The module H_0 becomes a $R_{\overline{\rho}}$-module via the map $R_{\overline{\rho}} \to \mathbb{T}_{U,\mathfrak{m}}$. We will in fact show something even stronger than $R_{\overline{\rho}} = \mathbb{T}_{U,\mathfrak{m}}$, namely that H_0 is a free $R_{\overline{\rho}}$-module.

The key input into the classical Taylor–Wiles method is the following additional data.

Proposition 3.1. *For every Taylor–Wiles set Q, we can find the following data:*

(i) *A finite free $W[\Delta_Q]$-module H_Q.*

(ii) *A homomorphism $R_{\overline{\rho},Q} \to \mathrm{End}_{W[\Delta_Q]}(H_Q)$ of $W[\Delta_Q]$-algebras.*

(iii) *An isomorphism $H_Q \otimes_{W[\Delta_Q]} W \cong H_0$ of W-modules such that the diagram*

$$\begin{array}{ccc} R_{\overline{\rho},Q} & \longrightarrow & \mathrm{End}_{W[\Delta_Q]}(H_Q) \\ \downarrow & & \downarrow{\scriptstyle -\otimes_{W[\Delta_Q]} W} \\ R_{\overline{\rho}} & \longrightarrow & \mathrm{End}_W(H_0) \end{array}$$

commutes.

Let $\Delta_\infty = \mathbb{Z}_p^r$, and let $S_\infty = W[\![\Delta_\infty]\!] \cong W[\![S_1, \ldots, S_r]\!]$. For every choice of Taylor–Wiles set Q, we fix a surjection $\Delta_\infty \to \Delta_Q$, and hence view $R_{\overline{\rho},Q}$ as S_∞-algebra. We also fix a surjection $W[\![X_1, \ldots, X_g]\!] \to R_{\overline{\rho},Q}$.

As Q varies, the kernels of the maps $\Delta_\infty \to \Delta_Q$ become arbitrarily small. A compactness (or "patching") argument (using the fact that all of the objects considered here are profinite) allows us to imagine that all of the data associated to Taylor–Wiles sets Q here are compatible as Q varies. More precisely, compactness allows us to pass to a convergent subsequence where this is indeed the case (see [Dia97, §2]). We can therefore 'pass to the limit with respect to Q' to obtain the following data:

(i) A finite free S_∞-module H_∞.

(ii) A complete local W-algebra $R_{\overline{\rho},\infty}$, together with a surjection $W[\![X_1, \ldots, X_g]\!] \to R_{\overline{\rho},\infty}$, and a homomorphism $S_\infty \to R_{\overline{\rho},\infty}$, together with an isomorphism $R_{\overline{\rho},\infty} \otimes_{S_\infty} W \cong R_{\overline{\rho}}$.

(iii) A homomorphism $R_{\overline{\rho},\infty} \to \mathrm{End}_{S_\infty}(H_\infty)$ of S_∞-algebras.

(iv) An isomorphism $H_\infty \otimes_{S_\infty} W \cong H_0$ of W-modules such that the diagram

$$\begin{array}{ccc} R_{\overline{\rho},\infty} & \longrightarrow & \mathrm{End}_{S_\infty}(H_\infty) \\ \downarrow & & \downarrow{\scriptstyle -\otimes_{S_\infty} W} \\ R_{\overline{\rho}} & \longrightarrow & \mathrm{End}_W(H_0) \end{array}$$

commutes.

[Warning: although elementary, the construction of these objects depends on many choices. In particular, there is no Galois representation with coefficients in the ring $R_{\overline{\rho},\infty}$.]

We have

$$\dim S_\infty = \mathrm{depth}_{S_\infty} H_\infty \leqslant \mathrm{depth}_{R_{\overline{\rho},\infty}} H_\infty \leqslant \dim R_{\overline{\rho},\infty} \leqslant \dim W[\![X_1, \dots, X_g]\!],$$

the first equality because H_∞ is free, the first inequality because the action of S_∞ on H_∞ factors through $R_{\overline{\rho},\infty}$, the second inequality because depth is bounded above by dimension, and the final inequality because $R_{\overline{\rho},\infty}$ is a quotient of the given power series ring.

At this point, the numerical coincidence $g = r$ intervenes to show that these inequalities are all equalities. This shows that the surjection $W[\![X_1, \dots, X_g]\!] \to R_{\overline{\rho},\infty}$ must be an isomorphism. To finish the argument, we use the Auslander–Buchsbaum formula $\mathrm{depth}_S M + \mathrm{proj\ dim}_S M = \dim S$, for a finitely generated module M over a regular local ring S. This implies that H_∞ is free over $R_{\overline{\rho},\infty}$, and hence $H_0 \cong H_\infty \otimes_{S_\infty} W$ is free over $R_{\overline{\rho}} \cong R_{\overline{\rho},\infty} \otimes_{S_\infty} W$. This is the end of the argument.

4. Beyond the classical case

We now return to the case of a general number field F, and take again $U = \prod_v \mathrm{GL}_2(\mathcal{O}_{F_v})$ and $\mathfrak{m} \subset \mathbb{T}_U$ to be a non-Eisenstein maximal ideal. Let $\overline{\rho} = \overline{\rho}_{\mathfrak{m}}$. Two problems arise when generalizing the arguments of the previous section:

- First, the numerical coincidence $g = r$ no longer applies, although it seemed to play a crucial role at the end of the argument. Instead, we have $g = r - r_2$, where r_2 is the number of complex places of F.
- Second, we still want to take $H_0 = H^*(X_U, W)_{\mathfrak{m}}$. However, the natural groups H_Q that you might write down are no longer free $W[\Delta_Q]$-modules. This freeness was important.

It turns out that these problems are connected, and will in the end cancel each other out. The second problem is closely related to the fact that the cohomology groups $H^*(X_U, W)_{\mathfrak{m}}$ are no longer concentrated in a single cohomological degree. This happens even rationally: if π is a cuspidal automorphic representation of $\mathrm{PGL}_2(\mathbb{A}_F)$ which contributes to $H^*(X_U, W)_{\mathfrak{m}}[1/p]$, then it will have non-zero contributions in all of the degrees in the range $[q_0, q_0 + l_0] = [r_1 + r_2, r_1 + 2r_2]$. (The notation q_0, l_0 introduced here generalizes away from GL_2, although the origin of these quantities is not important here.) Recall that conjecture B above states that the integral cohomology of $H^*(X_U, W)_{\mathfrak{m}}$ is concentrated in the same range.

We recall that we have defined the Hecke algebra $\mathbb{T}_U \subset \mathrm{End}_{\mathbf{D}(W)}(C(X_U, W))$. Since $\mathbb{T}_U$ is a finite W-algebra, and so p-adically complete, the maximal ideal $\mathfrak{m} \subset \mathbb{T}_U$ determines an idempotent $e_{\mathfrak{m}} \in \mathbb{T}_U$. We can then apply the following lemma to get a direct summand complex $C(X_U, W)_{\mathfrak{m}}$, which is equipped with a subalgebra $\mathbb{T}_{U,\mathfrak{m}} \subset \mathrm{End}_{\mathbf{D}(W)}(C(X_U, W)_{\mathfrak{m}})$ and a canonical isomorphism $H^*(C(X_U, W)_{\mathfrak{m}}) \cong H^*(X_U, W)_{\mathfrak{m}}$.

Lemma 4.1. *Let A be a ring and let $e \in \mathrm{End}_{\mathbf{D}(A)}(X)$ be an idempotent. Then there is an isomorphism $X \cong Y \oplus Z$ which identifies e with $\mathrm{id}_Y \oplus 0_Z$. Moreover, the objects Y, Z are uniquely determined up to isomorphism in $\mathbf{D}(A)$.*

In other words, the additive category $\mathbf{D}(A)$ is idempotent complete. We set $C_0 = C(X_U, W)_{\mathfrak{m}} \in \mathbf{D}(W)$.

To go further we now need to specify what additional structures there are given Taylor–Wiles data. We recall that given a Taylor–Wiles set $Q = \{v_1, \dots, v_r\}$, we have introduced the deformation ring $R_{\overline{\rho},Q}$, which parameterizes deformations of $\overline{\rho}$ which may now be ramified at places $v \in Q$: it is an algebra for the ring $W[\Delta_Q]$, where $\Delta_Q = \prod_{i=1}^r k(v_i)^\times(p)$.

Conjecture C. *For every Taylor–Wiles set Q, we can find the following data:*

(i) *A perfect complex C_Q of $W[\Delta_Q]$-modules.*

(ii) *A homomorphism $R_{\overline{\rho},Q} \to \mathrm{End}_{\mathbf{D}(W[\Delta_Q])}(C_Q)$ of $W[\Delta_Q]$-algebras.*

(iii) *An isomorphism $C_Q \otimes^{\mathbb{L}}_{W[\Delta_Q]} W \cong C_0$ in $\mathbf{D}(W)$ such that the diagram*

$$\begin{array}{ccc} R_{\overline{\rho},Q} & \longrightarrow & \mathrm{End}_{\mathbf{D}(W[\Delta_Q])}(C_Q) \\ \downarrow & & \downarrow{\scriptstyle -\otimes^{\mathbb{L}}_{W[\Delta_Q]} W} \\ R_{\overline{\rho}} & \longrightarrow & \mathrm{End}_{\mathbf{D}(W)}(C_0) \end{array}$$

commutes.

[The construction of C_Q will be explained in the next section; its definition is unconditional. The statement in (ii) that the map $R_{\overline{\rho},Q} \to \mathrm{End}_{\mathbf{D}(W[\Delta_Q])}(C_Q)$ respects $W[\Delta_Q]$-algebra structures is a slightly obscured form of local-global compatibility at Taylor–Wiles primes.]

Before we proceed to the patching argument, we recall a very useful lemma that explains the importance of Conjecture B. It is a version of Nakayama's lemma for complexes:

Lemma 4.2. *Let A be a Noetherian local ring.*

(i) *Say a complex $C \in \mathbf{D}(A)$ is minimal if C is a bounded complex of finite free A-modules and the differentials on $C \otimes_A A/\mathfrak{m}_A$ are 0. If C, D are minimal complexes of A-modules, and they are isomorphic in $\mathbf{D}(A)$, then this isomorphism is represented by an isomorphism $C \cong D$ in $\mathbf{K}(A)$.*

(ii) *Let C be a perfect complex of A-modules. Then C is isomorphic in $\mathbf{D}(A)$ to a minimal complex.*

If C is a minimal complex, then $C \otimes_A A/\mathfrak{m}_A = H^*(C \otimes_A A/\mathfrak{m}_A)$. Applying the traditional form of Nakayama's lemma, this shows that C^i is non-zero if and only if $H^i(C \otimes_A A/\mathfrak{m}_A) \neq 0$.

Applying the lemma (and Conjecture B) to C_0, we can assume that C_0 is a complex of finite free W-modules concentrated in the range $[q_0, q_0 + l_0]$. Applying the lemma to C_Q, we can assume that C_Q is a complex of finite free $W[\Delta_Q]$-modules concentrated in the range $[q_0, q_0 + l_0]$, and that the isomorphism of Conjecture C is represented by an isomorphism

$$C_Q \otimes_{W[\Delta_Q]} W \cong C_0$$

in $\mathbf{K}(W[\Delta_Q])$.

We can now carry out the Taylor–Wiles argument. Let $\Delta_\infty = \mathbb{Z}_p^r$, and let $S_\infty = W[\![\Delta_\infty]\!] \cong W[\![S_1, \dots, S_r]\!]$. We choose for every Taylor–Wiles set Q a surjection $\Delta_\infty \to \Delta_Q$ and a surjection

$W[\![X_1, \dots, X_g]\!] \to R_{\overline{\rho},Q}$. By another compactness argument (to 'take the limit $Q \to \infty$'; see §7 below), we obtain the following data:

(i) A complex C_∞ of finite free S_∞-modules, concentrated in degrees $[q_0, q_0 + l_0]$.

(ii) A complete local W-algebra $R_{\overline{\rho},\infty}$, together with a surjection $W[\![X_1, \dots, X_g]\!] \to R_{\overline{\rho},\infty}$, and a map $S_\infty \to R_{\overline{\rho},\infty}$, together with an isomorphism $R_{\overline{\rho},\infty} \otimes_{S_\infty} W \cong R_{\overline{\rho}}$.

(iii) A homomorphism $R_{\overline{\rho},\infty} \to \mathrm{End}_{\mathbf{D}(S_\infty)}(C_\infty)$ of S_∞-algebras.

(iv) An isomorphism $C_\infty \otimes_{S_\infty} W \cong C_0$ of complexes of W-modules such that the diagram

$$\begin{array}{ccc} R_{\overline{\rho},\infty} & \longrightarrow & \mathrm{End}_{\mathbf{D}(S_\infty)}(C_\infty) \\ \downarrow & & \downarrow{\scriptstyle -\otimes^{\mathbb{L}}_{S_\infty} W} \\ R_{\overline{\rho}} & \longrightarrow & \mathrm{End}_{\mathbf{D}(W)}(C_0) \end{array}$$

commutes.

We have

$$\dim_{S_\infty} H^*(C_\infty) \leqslant \dim_{R_{\overline{\rho},\infty}} H^*(C_\infty) \leqslant \dim R_{\overline{\rho},\infty} \leqslant \dim W[\![X_1, \dots, X_g]\!] = \dim S_\infty - l_0,$$

since the action of the ring S_∞ on the cohomology groups factors through $R_{\overline{\rho},\infty}$, and $r - g = l_0$. To conclude, we apply the following lemma of Calegari–Geraghty (see [CG, Lemma 6.2]):

Lemma 4.3. *Let S be a regular local ring of dimension d, let $0 \leqslant l \leqslant d$ be an integer, and let C be a complex of finite free S-modules, concentrated in degrees $[0, l]$. Then $\dim_S H^*(C) \geqslant d - l$, and if equality holds then there is a unique non-zero cohomology group $H^i(C)$ in degree $i = l$, and we have proj $\dim_S H^l(C) = l$, $\mathrm{depth}_S H^l(C) = d - l$.*

(Thus the complex C is a projective resolution of $H^l(C)$ of minimal length.) Applying this to the complex C_∞ of S_∞-modules shows that $\dim_{S_\infty} H^*(C) \geqslant \dim S_\infty - l_0$. We already know that the reverse inequality holds, so we deduce equality and can apply the lemma again to get

$$\begin{aligned} \dim S_\infty - l_0 = \mathrm{depth}_{S_\infty} H^{q_0+l_0}(C_\infty) &\leqslant \mathrm{depth}_{R_{\overline{\rho},\infty}} H^{q_0+l_0}(C_\infty) \\ &\leqslant \dim W[\![X_1, \dots, X_g]\!] = \dim S_\infty - l_0. \end{aligned}$$

Thus $\mathrm{depth}_{R_{\overline{\rho},\infty}} H^{q_0+l_0}(C_\infty) = \dim W[\![X_1, \dots, X_g]\!]$. We deduce that $R_{\overline{\rho},\infty} = W[\![X_1, \dots, X_g]\!]$ and, using the Auslander–Buchsbaum formula, that $H^{q_0+l_0}(C_\infty)$ is free over $R_{\overline{\rho},\infty}$. Finally we can pass to the quotient to deduce that $H^{q_0+l_0}(X_U, W)_{\mathfrak{m}} = H^{q_0+l_0}(C_\infty) \otimes_{S_\infty} W$ is free over $R_{\overline{\rho}} = R_{\overline{\rho},\infty} \otimes_{S_\infty} W$. In particular, the map $R_{\overline{\rho}} \to \mathbb{T}_{U,\mathfrak{m}}$ must be an isomorphism.

Looking at the proof, we see that it can be said to rely on a new numerical coincidence

$$\begin{array}{c} \dim_k H^1_{\mathcal{L}^\perp}(F_\Sigma/F, \mathrm{ad}^0\,\overline{\rho}(1)) \\ -\dim_k H^1_{\mathcal{L}}(F_\Sigma/F, \mathrm{ad}^0\,\overline{\rho}) \end{array} = l_0 = \begin{array}{c} \text{length of range in which} \\ \text{there is non-zero cohomology.} \end{array}$$

5. Construction of complexes

We now sketch the construction of the auxiliary complexes C_Q appearing in §4. Recall then that $C_0 = C(X_U, W)_{\mathfrak{m}}$, and we want to construct for every Taylor–Wiles set Q a perfect complex C_Q of $W[\Delta_Q]$-modules, together with a homomorphism $R_{\overline{\rho},Q} \to \mathrm{End}_{\mathbf{D}(W[\Delta_Q])}(C_Q)$ of $W[\Delta_Q]$-algebras

and an isomorphism $C_Q \otimes^{\mathbb{L}}_{W[\Delta_Q]} W \cong C_0$ such that the diagram

$$\begin{array}{ccc} R_{\overline{\rho},Q} & \longrightarrow & \mathrm{End}_{\mathbf{D}(W[\Delta_Q])}(C_Q) \\ \downarrow & & \downarrow{\scriptstyle -\otimes^{\mathbb{L}}_{W[\Delta_Q]} W} \\ R_{\overline{\rho}} & \longrightarrow & \mathrm{End}_{\mathbf{D}(W)}(C_0) \end{array}$$

commutes.

In the classical case $l_0 = 0$, the perfectness of the complex C_Q, together with Conjecture B, would imply that cohomology groups $H^*(C_Q)$ are free over $W[\Delta_Q]$ (and non-zero exactly in degree q_0). In general, this will definitely not be the case, and this necessitates the use of complexes rather than cohomology groups.

Let us therefore take a Taylor–Wiles set $Q = \{v_1, \dots, v_r\}$. We recall that by definition we have $\#k(v_i) = q_{v_i} \equiv 1 \bmod p$ for each $i = 1, \dots, r$, and the group Δ_Q is defined as the maximal p-power order quotient of the finite abelian group $\prod_{i=1}^r k(v_i)^\times$. We define open compact subgroups of $U = \prod_v \mathrm{GL}_2(\mathcal{O}_{F_v})$ as follows:

$$U_0(Q) = \prod_{v \notin Q} \mathrm{GL}_2(\mathcal{O}_{F_v}) \times \prod_{v \in Q} U_0(v),$$

$$U_1(Q) = \prod_{v \notin Q} \mathrm{GL}_2(\mathcal{O}_{F_v}) \times \prod_{v \in Q} U_1(v),$$

where

$$U_0(v) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathcal{O}_{F_v}) \mid c \equiv 0 \bmod (\varpi_v) \right\},$$

$$U_1(v) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in U_0(v) \mid ad^{-1} \bmod (\varpi_v) \text{ has order prime to } p \right\}.$$

Then $U_0(Q) \subset U$ has finite index prime to p (its index is $\prod_{i=1}^r (q_{v_i} + 1)$ and p is odd), and $U_1(Q) \subset U_0(Q)$ is a normal subgroup with $U_0(Q)/U_1(Q) \cong \Delta_Q$. Since we assume that U is neat, the natural map $X_{U_1(Q)} \to X_{U_0(Q)}$ is a Galois covering, with group Δ_Q. Then we can upgrade $C(U_1(Q), W)$ to be a perfect complex of $W[\Delta_Q]$-modules such that

$$H^*(C(U_1(Q), W)) = H^*(X_{U_1(Q)}, W)$$

and

$$H^*(C(U_1(Q), W) \otimes^{\mathbb{L}}_{W[\Delta_Q]} W) = H^*(X_{U_0(Q)}, W).$$

The Hecke algebra $\mathbb{T}_U$ lives in $\mathrm{End}_{\mathbf{D}(W)}(C(X_U, W))$, while the Hecke algebra $\mathbb{T}_{U_0(Q)}$ lives in $\mathrm{End}_{\mathbf{D}(W)}(C(X_{U_0(Q)}, W)$. We define $\mathbb{T}_{U_0(Q)/U_1(Q)}$ to be the $W[\Delta_Q]$-subalgebra of

$$\mathrm{End}_{\mathbf{D}(W[\Delta_Q])}(C(X_{U_1(Q)}, W))$$

generated by unramified Hecke operators. Then there is a commutative diagram

$$\begin{array}{ccc} \mathbb{T}_{U_0(Q)/U_1(Q)} & \longrightarrow & \mathrm{End}_{\mathbf{D}(W[\Delta_Q])}(C(X_{U_1(Q)}, W)) \\ \downarrow & & \downarrow{\scriptstyle -\otimes^{\mathbb{L}}_{W[\Delta_Q]} W} \\ \mathbb{T}_{U_0(Q)} & \longrightarrow & \mathrm{End}_{\mathbf{D}(W)}(C(X_{U_0(Q)}, W)). \end{array}$$

The index $[U : U_0(Q)] = \prod_{i=1}^r (q_{v_i} + 1)$ is prime to p, so the pullback morphism $C(X_U, W) \to$

$C(X_{U_0(Q)}, W)$ is split by the trace map $\mathrm{tr}_{U/U_0(Q)}$. This allows us to write down another commutative diagram

$$\begin{array}{ccc} \mathbb{T}_{U_0(Q)} & \longrightarrow & \mathrm{End}_{\mathbf{D}(W)}(C(X_{U_0(Q)}, W)) \\ \downarrow & & \downarrow \\ \mathbb{T}_U & \longrightarrow & \mathrm{End}_{\mathbf{D}(W)}(C(X_U, W)). \end{array}$$

We write $\mathfrak{n}_0$ for the pullback of $\mathfrak{m}$ to $\mathbb{T}_{U_0(Q)}$ and $\mathfrak{n}_1$ for the pullback of $\mathfrak{n}_0$ to $\mathbb{T}_{U_0(Q)/U_1(Q)}$. Then there are surjective maps

$$\mathbb{T}_{U_0(Q)/U_1(Q),\mathfrak{n}_1} \to \mathbb{T}_{U_0(Q),\mathfrak{n}_0} \to \mathbb{T}_{U,\mathfrak{m}}$$

and we can define the corresponding localized complexes

$$C(X_{U_1(Q)}, W)_{\mathfrak{n}_1} \in \mathbf{D}(W[\Delta_Q])$$

and

$$C(X_{U_0(Q)}, W)_{\mathfrak{n}_0} \in \mathbf{D}(W).$$

Then there is an isomorphism

$$C(X_{U_1(Q)}, W)_{\mathfrak{n}_1} \otimes^{\mathbb{L}}_{W[\Delta_Q]} W \cong C(X_{U_0(Q)}, W)_{\mathfrak{n}_0}.$$

A souped-up version of Conjecture A would imply the existence of a lifting of $\overline{\rho}_{\mathfrak{m}}$ to $\mathbb{T}_{U_0(Q)/U_1(Q),\mathfrak{n}_1}$, hence a surjective map $R_{\overline{\rho},Q} \to \mathbb{T}_{U_0(Q)/U_1(Q),\mathfrak{n}_1}$. We are still not quite finished: even in the classical case, this map will not be a homomorphism of $W[\Delta_Q]$-algebras (which is a key requirement).

To get around this, we introduce the double coset operators $U_v = [U_1(v)\,\mathrm{diag}(\varpi_v, 1)U_1(v)]$ for $v \in Q$ and define the enlarged Hecke algebras

$$\mathbb{T}^{Q\text{-aug}}_{U_0(Q)/U_1(Q),\mathfrak{n}_1} \subset \mathrm{End}_{\mathbf{D}(W[\Delta_Q])}(C(X_{U_1(Q)}, W)_{\mathfrak{n}_1})$$

and

$$\mathbb{T}^{Q\text{-aug}}_{U_0(Q),\mathfrak{n}_0} \subset \mathrm{End}_{\mathbf{D}(W)}(C(X_{U_0(Q)}, W)_{\mathfrak{n}_0})$$

to be the algebras generated by $\mathbb{T}_{U_0(Q)/U_1(Q),\mathfrak{n}_1}$ and $\mathbb{T}_{U_0(Q),\mathfrak{n}_0}$, respectively, and all the operators U_v ($v \in Q$). A calculation in the local Iwahori–Hecke algebra $\mathcal{H}(\mathrm{GL}_2(F_v), U_0(Q))$ (see [KT, §5], where this calculation is done for GL_n; a simpler argument suffices when $n = 2$, see [CG]) shows that the ideals

$$\mathfrak{n}_{0,\alpha} = (\mathfrak{n}_0, U_{v_1} - \alpha_1, \ldots, U_{v_r} - \alpha_r) \subset \mathbb{T}^{Q\text{-aug}}_{U_0(Q),\mathfrak{n}_0}$$

and

$$\mathfrak{n}_{1,\alpha} = (\mathfrak{n}_1, U_{v_1} - \alpha_1, \ldots, U_{v_r} - \alpha_r) \subset \mathbb{T}^{Q\text{-aug}}_{U_0(Q)/U_1(Q),\mathfrak{n}_1}$$

are proper maximal ideals which appear in the support of $H^*(X_{U_0(Q)}, k)$ and $H^*(X_{U_1(Q)}, k)$, respectively. (We recall that for each $v_i \in Q$, $\overline{\rho}(\mathrm{Frob}_v)$ is assumed to have distinct eigenvalues $\alpha_i, \beta_i \in k$.) We consider the further localizations

$$C(X_{U_1(Q)}, W)_{\mathfrak{n}_{1,\alpha}},$$

direct summand of $C(X_{U_1(Q)}, W)_{\mathfrak{n}_1}$, and

$$C(X_{U_0(Q)}, W)_{\mathfrak{n}_{0,\alpha}},$$

direct summand of $C(X_{U_0(Q)}, W)_{\mathfrak{n}_0}$. These complexes have the property that

$$C(X_{U_1(Q)}, W)_{\mathfrak{n}_{1,\alpha}} \otimes^{\mathbb{L}}_{W[\Delta_Q]} W \cong C(X_{U_0(Q)}, W)_{\mathfrak{n}_{0,\alpha}}$$

and one can moreover show (cf. [KT, §5]) that there is now an isomorphism in $\mathbf{D}(W)$

$$C(X_{U_0(Q)}, W)_{\mathfrak{n}_{0,\alpha}} \cong C(X_U, W)_{\mathfrak{m}}$$

which respects the action of unramified Hecke operators. We can finally therefore take

$$C_Q = C(X_{U_1(Q)}, W)_{\mathfrak{n}_{1,\alpha}},$$

with its induced map $R_{\overline{\rho},Q} \to \mathrm{End}_{\mathbf{D}(W[\Delta_Q])}(C_Q)$. To verify Conjecture C, it remains to show that this map is a homomorphism of $W[\Delta_Q]$-algebras. This would follow from an appropriate version of local-global compatibility.

In the classical case, the above complexes are isomorphic in $\mathbf{D}(W[\Delta_Q])$ to their unique non-zero cohomology groups, and our construction reduces to the one given in e.g. [Gee, §5.5].

6. Further remarks

There is much more to say than we have time for here:

Remark. (i) Everything we have said generalizes essentially without change to the case of GL_n (in particular, there is no need to fix the determinant), as long as one states the correct versions of the corresponding conjectures. The same ideas also apply in other situations where the numerical coincidence fails to hold. A key example is the case of weight 1 modular forms for GL_2 over $\mathbb{Q}$ (where there are repeated Hodge–Tate weights). Calegari–Geraghty implement their ideas in [CG] to prove unconditional $R = \mathbb{T}$ theorems for weight 1 modular forms.

(ii) We have shown above how to prove an $R = \mathbb{T}$ theorem. This leads to a traditional automorphy lifting theorem of the following form:

Theorem 6.1. *Let p be an odd prime and let F be a number field unramified at p. Assume Conjectures A, B and C above. Let $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$ be a continuous representation, unramified away from p and finite flat at p, and such that $\det \rho = \epsilon^{-1}$. Suppose that there exists a cuspidal automorphic representation π of $\mathrm{PGL}_2(\mathbb{A}_F)$, cohomological of weight 0 and everywhere unramified, such that $\overline{r_\iota(\pi)} \cong \overline{\rho}$ (for some choice of isomorphism $\iota : \overline{\mathbb{Q}}_p \cong \mathbb{C}$). Suppose moreover that $\overline{\rho}$ satisfies a big image hypotheses (e.g. $p > 5$ and $\overline{\rho}|_{G_{F(\zeta_p)}}$ absolutely irreducible). Then ρ is automorphic: there exists a cuspidal automorphic representation Π of $\mathrm{PGL}_2(\mathbb{A}_F)$, cohomological of weight 0, and an isomorphism $\rho \cong r_\iota(\Pi)$.*

It is obviously desirable to have a version of this theorem that allows ρ and π to be ramified (at least at the primes away from p). In [CG] it is shown that the technique used in [Tay08] can be generalized to the current setting. This allows one to prove automorphy lifting results where ρ and π can have arbitrary ramification away from p (still assuming the appropriate generalizations of Conjectures A, B and C). It is shown in [CG] that these (conditional) theorems are strong enough to imply e.g. the potential modularity of all elliptic curves over all number fields!

(iii) The Galois representations that are required in Conjecture A have been almost shown to exist by Scholze in the case where F is a CM field [Sch15]. More precisely he constructs representations with coefficients in a quotient $\mathbb{T}_U/I$, where I is an ideal satisfying $I^\delta = 0$ for some explicit integer δ (depending only on F and n, and not on the particular choice of U). One can modify the arguments here to find that this weaker statement is still enough to prove results like Theorem 6.1, although one will no longer be able to show that the map $R_{\overline{\rho}} \to \mathbb{T}_{U,\mathfrak{m}}$ is an isomorphism on the nose.

The modifications required in order to deal with the presence of a nilpotent ideal I are quite simple, and there is some hope (based on [NT16]) that one will eventually have access to Scholze's results with $\delta = 1$ (i.e. $I = 0$), so we have avoided discussing them in the main part of these notes. We discuss the modifications required in §8 below.

7. Appendix 1: how to patch

Here we describe an abstract result that describes how to patch complexes together. The result and its proof are based on [KT, Proposition 3.1]. We still use k to denote a finite field and W its ring of Witt vectors.

Proposition 7.1. *Fix integers $g, r \geqslant 0$, and let $\Delta_\infty = \mathbb{Z}_p^r$. If $N \geqslant 0$, let $\Delta_N = \Delta_\infty/(p^N)$. Let $S_\infty = W[\![\Delta_\infty]\!]$. Suppose given the following data:*

(i) *A complete Noetherian local W-algebra R_0 and a surjection $W[\![X_1, \dots, X_g]\!] \to R_0$ of W-algebras.*

(ii) *A minimal complex C_0 of W-modules.*

(iii) *A homomorphism $R_0 \to \mathrm{End}_{\mathbf{D}(W)}(C_0)$ of W-algebras.*

Suppose given as well for every $N \geqslant 1$ the following data:

(i) *A complete Noetherian local $W[\Delta_N]$-algebra R_N and a surjection $W[\![X_1, \dots, X_g]\!] \to R_N$ of W-algebras.*

(ii) *A minimal complex C_N of $W[\Delta_Q]$-modules.*

(iii) *A homomorphism $R_N \to \mathrm{End}_{\mathbf{D}(W[\Delta_N])}(C_N)$ of $W[\Delta_N]$-algebras.*

(iv) *Isomorphisms $R_N \otimes_{W[\Delta_N]} W \cong R_0$ and $C_N \otimes_{W[\Delta_N]} W \cong C_0$ such that the diagram*

$$\begin{array}{ccc} R_N & \longrightarrow & \mathrm{End}_{\mathbf{D}(W[\Delta_N])}(C_N) \\ \downarrow & & \downarrow{\scriptstyle -\otimes^{\mathbb{L}}_{W[\Delta_N]} W} \\ R_0 & \longrightarrow & \mathrm{End}_{\mathbf{D}(W)}(C_0) \end{array}$$

commutes.

Then we can find the following data:

(i) *A complete Noetherian local S_∞-algebra R_∞ and a surjection $W[\![X_1, \dots, X_g]\!] \to R_\infty$ of W-algebras.*

(ii) *A minimal complex C_∞ of S_∞-modules.*

(iii) *A homomorphism $R_\infty \to \mathrm{End}_{\mathbf{D}(S_\infty)}(C_\infty)$ of S_∞-algebras.*

(iv) *Isomorphisms $R_\infty \otimes_{S_\infty} W \cong R_0$ and $C_\infty \otimes_{S_\infty} W \cong C_0$ such that the diagram*

$$\begin{array}{ccc} R_\infty & \longrightarrow & \mathrm{End}_{\mathbf{D}(S_\infty)}(C_\infty) \\ \downarrow & & \downarrow{\scriptstyle -\otimes^{\mathbb{L}}_{S_\infty} W} \\ R_0 & \longrightarrow & \mathrm{End}_{\mathbf{D}(W)}(C_0) \end{array}$$

commutes.

Proof. Let $s = \dim_k H^*(C_0 \otimes_W k)$, and suppose that C_0 is concentrated in the range $[0, d]$ for some $d \geqslant 0$. If $t \in \mathfrak{m}_{R_N}$, then t acts nilpotently on $H^*(C_N \otimes_{W[\Delta_N]} k) = H^*(C_0 \otimes_W k)$, so t^s acts

as 0 on $H^*(C_N \otimes_{W[\Delta_N]} k)$. The existence of the spectral sequence of a filtered complex implies that t^{N^2rs} acts as 0 on $H^*(C_N \otimes_{W[\Delta_N]} W[\Delta_N]/(p^N))$, and then [KT, Lemma 2.5] shows that $t^{(d+1)N^2rs}$ has trivial image in

$$\mathrm{End}_{\mathbf{D}(W[\Delta_N]/(p^N))}(C_N \otimes_{W[\Delta_N]} W[\Delta_N]/(p^N)).$$

It follows that the map

$$R_N \to \mathrm{End}_{\mathbf{D}(W[\Delta_N]/(p^N))}(C_N \otimes_{W[\Delta_N]} W[\Delta_N]/(p^N))$$

factors through the quotient $R_N/\mathfrak{m}_{R_N}^{(g+1)(d+1)N^2rs}$.

We define a patching datum of level $N \geqslant 1$ to be a tuple $(D, \psi, R, \eta_0, \eta_1, \eta_2)$, where:

- D is a minimal complex of $W[\Delta_N]/(p^N)$-modules and ψ is an isomorphism
$$\psi : D \otimes_{W[\Delta_N]/(p^N)} W/(p^N) \cong C_0 \otimes_W W/(p^N).$$
- R is a complete Noetherian local S_∞-algebra equipped with a surjection
$$\eta_0 : W[\![X_1, \dots, X_g]\!] \to R$$
of W-algebras, a homomorphism
$$\eta_1 : R \to \mathrm{End}_{\mathbf{D}(W[\Delta_N]/(p^N))}(D)$$
of S_∞-algebras, and a homomorphism
$$\eta_2 : R \to R_0/\mathfrak{m}_{R_0}^{(g+1)(d+1)N^2rs}$$
of S_∞-algebras. Moreover, $\mathfrak{m}_R^{(g+1)(d+1)N^2rs} = 0$, and the diagram
$$\begin{array}{ccc} R & \longrightarrow & \mathrm{End}_{\mathbf{D}(W[\Delta_N]/(p^N))}(D) \\ \downarrow & & \downarrow \\ R_0 & \longrightarrow & \mathrm{End}_{\mathbf{D}(W/(p^N))}(C_0) \end{array}$$
commutes.

We make the collection of patching data of level N into a category Patch_N as follows: a morphism $\alpha : (D, \psi, R, \eta_0, \eta_1, \eta_2) \to (D', \psi', R', \eta_0', \eta_1', \eta_2')$ is a pair $\alpha = (f, g)$, where:

- $f : D \to D'$ is an isomorphism of complexes of $W[\Delta_N]/(p^N)$-modules such that $\psi'(f \otimes_{W[\Delta_N]/(p^N)} W/(p^N))\psi^{-1}$ is the identity.
- $g : R \to R'$ is an isomorphism of S_∞-algebras that intertwines η_0 and η_0', η_1 and η_1', and η_2 and η_2'.

Evidently the category Patch_N is a groupoid, i.e. every morphism is an isomorphism. There is a collection of functors $(F_N : \mathrm{Patch}_{N+1} \to \mathrm{Patch}_N)_{N \geqslant 1}$, which assign to a tuple $(D, \psi, R, \eta_0, \eta_1, \eta_2)$ the tuple $F_N(D, \psi, R, \eta_0, \eta_1, \eta_2) = (D', \psi', R', \eta_0', \eta_1', \eta_2')$ given as follows:

- $D' = D \otimes_{W[\Delta_{N+1}]/(p^{N+1})} W[\Delta_N]/(p^N)$.
- ψ' is the composite
$$D' \otimes_{W[\Delta_N]/(p^N)} W/(p^N) \cong D \otimes_{W[\Delta_{N+1}]/(p^{N+1})} W/(p^N) \cong C_0 \otimes_W W/(p^N).$$
- $R' = R/\mathfrak{m}_R^{(g+1)(d+1)N^2rs}$.

– η_0', η_1' and η_2' are the obvious maps.

The set of isomorphism classes of patching data of level N is finite. Indeed, it suffices to note that the cardinality of the ring R in the tuple $(D, \psi, R, \eta_0, \eta_1, \eta_2)$ is bounded solely in terms of N (since $W[\![X_1, \dots, X_g]\!]/(p, X_1, \dots, X_g)^{(g+1)(d+1)N^2rs}$ is a ring of finite cardinality), and $\mathrm{End}_{\mathbf{D}(W[\Delta_N]/(p^N))}(D)$ has cardinality bounded above by $|W[\Delta_N]/(p^N)| \cdot (\dim_k D \otimes_{W[\Delta_N]/(p^N)} k)^2$. (Indeed, since D is a complex of finite free modules, every endomorphism of D in $\mathbf{D}(W[\Delta_N]/(p^N))$ is represented by a homotopy class of endomorphisms of D as a complex.)

The set of isomorphism classes is also non-empty. Indeed, for each $M \geqslant 1$, we can define a patching datum $D(M, M) = (D, \psi, R, \eta_0, \eta_1, \eta_2) \in \mathrm{Patch}_M$ as follows:

– $D = C_M \otimes_{W[\Delta_M]} W[\Delta_M]/(p^M)$.

– ψ is the reduction modulo p^M of the given isomorphism $C_M \otimes_{W[\Delta_M]} W \cong C_0$.

– $R = R_M/\mathfrak{m}_{R_M}^{(g+1)(d+1)M^2rs}$.

– η_0, η_1 and η_2 are the obvious maps.

If $M \geqslant N \geqslant 1$, then we define $D(M, N) \in \mathrm{Patch}_N$ to be the image of $D(M, M)$ under the composite functor $F_N \circ F_{N+1} \circ \cdots \circ F_{M-1}$. After diagonalization, we can find an increasing sequence $(M_N)_{N \geqslant 1}$ of integers, together with a system of isomorphisms $\alpha_N : F_N(D(M_{N+1}, N+1)) \to D(M_N, N)$. Passing to the inverse limit with respect to these isomorphisms, we obtain a tuple $(C_\infty, \psi_\infty, R_\infty, \eta_0, \eta_1, \eta_2)$, where:

– C_∞ is a bounded complex of finite free S_∞-modules and ψ_∞ is an isomorphism

$$\psi_\infty : C_\infty \otimes_{S_\infty} W \cong C_0.$$

– R_∞ is a complete Noetherian local S_∞-algebra equipped with a surjection

$$\eta_0 : W[\![X_1, \dots, X_g]\!] \to R_\infty$$

of W-algebras, a homomorphism

$$\eta_1 : R_\infty \to \mathrm{End}_{\mathbf{D}(S_\infty)}(C_\infty)$$

of S_∞-algebras, and a homomorphism

$$\eta_2 : R_\infty \to R_0$$

of S_∞-algebras. Moreover, the diagram

$$\begin{array}{ccc} R_\infty & \longrightarrow & \mathrm{End}_{\mathbf{D}(S_\infty)}(C_\infty) \\ \downarrow & & \downarrow \\ R_0 & \longrightarrow & \mathrm{End}_{\mathbf{D}(W)}(C_0) \end{array}$$

commutes.

(We have used here the fact that if F is a bounded complex of finite free S_∞-modules, then there is a canonical isomorphism

$$\mathrm{End}_{\mathbf{D}(S_\infty)}(F) = \varprojlim \mathrm{End}_{\mathbf{D}(W[\Delta_N]/(p^N))}(F \otimes_{S_\infty} W[\Delta_N]/(p^N)).$$

This is a consequence of the representation of endomorphisms in $\mathbf{D}$ by homotopy classes of endomorphisms as complexes, and a Mittag-Leffler argument, see the proof of [KT, Lemma 2.13].) This completes the proof. □

8. Appendix 2: nilpotent ideals

In this section we briefly discuss how to modify the arguments of §4 assuming weaker versions of Conjectures A and C. We continue to assume that F is a number field and that p is an odd prime unramified in F, and now state these weaker versions:

Conjecture A′. *Let $U = \prod_v U_v \subset \mathrm{GL}_2(\widehat{\mathcal{O}}_F)$ be an open subgroup, and let $S \supset \Sigma$ be a finite set of finite places such that $U_v = \mathrm{GL}_2(\mathcal{O}_{F_v})$ if $v \notin S$.*

(i) *For each maximal ideal $\mathfrak{m} \subset \mathbb{T}_U$, there exists a continuous representation $\overline{\rho}_{\mathfrak{m}} : \mathrm{Gal}(F_S/F) \to \mathrm{GL}_2(\mathbb{T}_U/\mathfrak{m})$ such that $\det \overline{\rho}_{\mathfrak{m}} = \epsilon^{-1}$ and for all $v \notin S$, $\mathrm{tr}\, \overline{\rho}_{\mathfrak{m}}(\mathrm{Frob}_v) = T_v \bmod \mathfrak{m}$.*

(ii) *If $\mathfrak{m} \subset \mathbb{T}_U$ is a maximal ideal such that $\overline{\rho}_{\mathfrak{m}}$ is absolutely irreducible (in which case we say that $\mathfrak{m}$ is non-Eisenstein), then there exists an absolute constant $\delta = \delta(F)$, an ideal $I \subset \mathbb{T}_{U,\mathfrak{m}}$ such that $I^\delta = 0$, and a lifting*

$$\rho_{\mathfrak{m}} : \mathrm{Gal}(F_S/F) \to \mathrm{GL}_2(\mathbb{T}_{U,\mathfrak{m}}/I)$$

of $\overline{\rho}_{\mathfrak{m}}$ such that $\det \rho_{\mathfrak{m}} = \epsilon^{-1}$ and for all $v \notin S$, $\mathrm{tr}\, \rho_{\mathfrak{m}}(\mathrm{Frob}_v) = T_v \bmod I$.

(iii) *If $\mathfrak{m}$ is a non-Eisenstein maximal ideal and $U_v = \mathrm{GL}_2(\mathcal{O}_{F_v})$ for all $v \in \Sigma$, then the representation $\rho_{\mathfrak{m}}$ is finite flat at p.*

If F is an imaginary CM field, then parts (i) and (ii) of Conjecture A′ follow from the work of Scholze [Sch15]. It is proved in [NT16] that in this case, one can even take $\delta = 4$.

Let $U = \prod_v \mathrm{GL}_2(\mathcal{O}_{F_v})$, and let $\mathfrak{m} \subset \mathbb{T}_U$ be a non-Eisenstein maximal ideal. Let $\overline{\rho} = \overline{\rho}_{\mathfrak{m}}$. Let $\mathbb{T}_0 = \mathbb{T}_{U,\mathfrak{m}}$, $I_0 \subset \mathbb{T}_0$ the ideal given by Conjecture A′. We recall that $C_0 = C(U, W)_{\mathfrak{m}}$.

Conjecture C′. *For every Taylor–Wiles set Q, we can find the following data:*

(i) *A perfect complex C_Q of $W[\Delta_Q]$-modules and an isomorphism $C_Q \otimes^{\mathbb{L}}_{W[\Delta_Q]} W \cong C_0$ in $\mathbf{D}(W)$.*

(ii) *A $W[\Delta_Q]$-subalgebra $\mathbb{T}_Q \subset \mathrm{End}_{\mathbf{D}(W[\Delta_Q])}(C_Q)$ such that the image of $\mathbb{T}_Q$ under the map $\mathrm{End}_{\mathbf{D}(W[\Delta_Q])}(C_Q) \to \mathrm{End}_{\mathbf{D}(W)}(C_0)$ equals $\mathbb{T}_0$.*

(iii) *An ideal $I_Q \subset \mathbb{T}_Q$ such that $I_Q^\delta = 0$ and a surjective homomorphism $R_{\overline{\rho},Q} \to \mathbb{T}_Q/I_Q$ of $W[\Delta_Q]$-algebras such that the diagram*

$$\begin{array}{ccccccc} R_{\overline{\rho},Q} & \longrightarrow & \mathbb{T}_Q/I_Q & \longleftarrow & \mathbb{T}_Q & \longrightarrow & \mathrm{End}_{\mathbf{D}(W[\Delta_Q])}(C_Q) \\ \downarrow & & \downarrow & & \downarrow & & \downarrow{\scriptstyle -\otimes^{\mathbb{L}}_{W[\Delta_Q]} W} \\ R_{\overline{\rho}} & \longrightarrow & \mathbb{T}_0/(I_0, I_Q) & \longleftarrow & \mathbb{T}_0 & \longrightarrow & \mathrm{End}_{\mathbf{D}(W)}(C_0) \end{array}$$

commutes.

Note that Conjecture C′ with $\delta = 1$ is equivalent to Conjecture C. We now assume Conjectures A′, B, and C′, and will show that the surjective map $R_{\overline{\rho}} \to \mathbb{T}_0/I_0$ is an isomorphism after passage to reduced quotients. In particular, this statement suffices to prove an automorphy lifting theorem (for $\overline{\mathbb{Q}}_p$-valued lifts of $\overline{\rho}$).

As in §4, we assume that C_0 is represented by a minimal complex of W-modules, hence concentrated in the range $[q_0, q_0 + l_0] = [r_1 + r_2, r_1 + 2r_2]$, and similarly for all the complexes C_Q. Using a simple variant of Proposition 7.1, we obtain the following objects (with $S_\infty = W[\![\mathbb{Z}_p^r]\!]$):

(i) A complex C_∞ of finite free S_∞-modules and an isomorphism $C_\infty \otimes_{S_\infty} W \cong C_0$ in $\mathbf{K}(W)$.

(ii) An S_∞-subalgebra $\mathbb{T}_\infty \subset \mathrm{End}_{\mathbf{D}(S_\infty)}(C_\infty)$ such that the image of $\mathbb{T}_\infty$ under the map $\mathrm{End}_{\mathbf{D}(S_\infty)}(C_\infty) \to \mathrm{End}_{\mathbf{D}(W)}(C_0)$ equals $\mathbb{T}_0$.

(iii) An ideal $I_\infty \subset \mathbb{T}_\infty$ such that $I_\infty^\delta = 0$, a complete Noetherian local S_∞-algebra $R_{\overline{\rho},\infty}$, a surjection $W[\![X_1, \dots, X_g]\!] \to R_{\overline{\rho},\infty}$ of W-algebras, an isomorphism $R_{\overline{\rho},\infty} \otimes_{S_\infty} W \cong R_{\overline{\rho}}$, and a surjective map $R_{\overline{\rho},\infty} \to \mathbb{T}_\infty/I_\infty$ of S_∞-algebras such that the diagram

$$\begin{array}{ccccccc} R_{\overline{\rho},\infty} & \longrightarrow & \mathbb{T}_\infty/I_\infty & \longleftarrow & \mathbb{T}_\infty & \longrightarrow & \mathrm{End}_{\mathbf{D}(S_\infty)}(C_\infty) \\ \downarrow & & \downarrow & & \downarrow & & \downarrow{\scriptstyle -\otimes^{\mathbb{L}}_{S_\infty} W} \\ R_{\overline{\rho}} & \longrightarrow & \mathbb{T}_0/(I_0, I_\infty) & \longleftarrow & \mathbb{T}_0 & \longrightarrow & \mathrm{End}_{\mathbf{D}(W)}(C_0) \end{array}$$

commutes.

We can now conclude the argument along similar lines to before. Conjecture B and Lemma 4.3 show that $\dim_{S_\infty} H^*(C_\infty) \geqslant \dim S_\infty - l_0$. On the other hand, we have

$$\dim_{S_\infty} H^*(C_\infty) = \dim_{\mathbb{T}_\infty} H^*(C_\infty) \leqslant \dim \mathbb{T}_\infty = \dim \mathbb{T}_\infty/I_\infty \leqslant \dim R_{\overline{\rho},\infty} \leqslant W[\![X_1, \dots, X_g]\!]$$

(since quotient by nilpotent ideals does not change dimension). We now use the numerical coincidence to find that all these inequalities are equalities, and apply Lemma 4.3 once more to conclude that $H^*(C_\infty) = H^{q_0+l_0}(C_\infty)$ is an S_∞-module of depth $\dim S_\infty - l_0$, hence a $\mathbb{T}_\infty$-module of depth $\dim S_\infty - l_0$. We also see that the maps $W[\![X_1, \dots, X_g]\!] \to R_{\overline{\rho},\infty}$ and $R_{\overline{\rho},\infty} \to \mathbb{T}_\infty/I_\infty$ are isomorphisms.

We can now apply e.g. [Tay08, Lemma 2.3] to conclude that $H^*(C_\infty)$ is a nearly faithful $\mathbb{T}_\infty$-module; by definition, this means that the ideal $\mathrm{Ann}_{\mathbb{T}_\infty} H^*(C_\infty)$ is nilpotent. We now repeatedly apply [Tay08, Lemma 2.2] (which says that if A is Noetherian local with ideal $I \subset A$ and M is a nearly faithful finite A-module, then $M/(I)$ is nearly faithful over A/I). First, we see that $H^*(C_\infty)/(I_\infty)$ is a nearly faithful $\mathbb{T}_\infty/I_\infty = R_{\overline{\rho},\infty}$-module. Let $\mathfrak{a} = \ker(S_\infty \to W)$; then we see that $H^{q_0+l_0}(C_\infty)/(I_\infty, \mathfrak{a}) = H^{q_0+l_0}(C_0)/(I_\infty)$ is a nearly faithful $R_{\overline{\rho},\infty}/(\mathfrak{a}) = R_{\overline{\rho}}$-module. This completes the proof.

References

CG Frank Calegari and David Geraghty. Modularity lifting beyond the Taylor–Wiles method. Preprint.

CHT08 Laurent Clozel, Michael Harris, and Richard Taylor. Automorphy for some l-adic lifts of automorphic mod l Galois representations. *Publ. Math. Inst. Hautes Études Sci.*, (108):1–181, 2008. With Appendix A, summarizing unpublished work of Russ Mann, and Appendix B by Marie-France Vignéras.

DDT97 Henri Darmon, Fred Diamond, and Richard Taylor. Fermat's last theorem. In *Elliptic curves, modular forms & Fermat's last theorem (Hong Kong, 1993)*, pages 2–140. Int. Press, Cambridge, MA, 1997.

Dia97 Fred Diamond. The Taylor-Wiles construction and multiplicity one. *Invent. Math.*, 128(2):379–391, 1997.

Dim05 Mladen Dimitrov. Galois representations modulo p and cohomology of Hilbert modular varieties. *Ann. Sci. École Norm. Sup. (4)*, 38(4):505–551, 2005.

Gee Toby Gee. Notes from a course at the Arizona Winter School, 2013. Available at `http://www2.imperial.ac.uk/~tsg/Index_files/ArizonaWinterSchool2013.pdf`.

Han David Hansen. Minimal modularity lifting for $GL(2)$ over an arbitrary number field. To appear in Math. Res. Lett.

KT Chandrashekhar Khare and Jack A. Thorne. Potential automorphy and the Leopoldt conjecture. To appear in AJM.

NT16 James Newton and Jack A. Thorne. Torsion Galois representations over CM fields and Hecke algebras in the derived category. *Forum Math. Sigma*, 4:e21, 88, 2016.

Sch15 Peter Scholze. On torsion in the cohomology of locally symmetric varieties. *Ann. of Math. (2)*, 182(3):945–1066, 2015.

Tay08 Richard Taylor. Automorphy for some l-adic lifts of automorphic mod l Galois representations. II. *Publ. Math. Inst. Hautes Études Sci.*, (108):183–239, 2008.

TW95 Richard Taylor and Andrew Wiles. Ring-theoretic properties of certain Hecke algebras. *Ann. of Math. (2)*, 141(3):553–572, 1995.