PATCHING AND MODULARITY

Contents

1. Introduction

This is the third part of the notes on modularity theorems. The aim of the third part is to study modularity/ automorphy lifting theorems for two-dimensional p-adic representations, using wherever possible arguments that go over to the n-dimensional case. The main tool is the Taylor-Wiles-Kisin patching method. To demonstrate the power of this method, we prove a modularity result following the strategy in [36]. Let $p > 3$ be a prime number. Let $L/\mathbb{Q}_p$ be a finite extension with ring of integers $\mathcal{O}$, maximal ideal λ, residue field $\mathbb{F} = \mathcal{O}/\lambda$. Let F be a totally real field. Assume that L is sufficiently large so that L contains the images of all embeddings $F \hookrightarrow \overline{L}$.

Theorem 1.1. *Let ρ, $\rho_0 : G_F \to \mathrm{GL}_2(\mathcal{O})$ be two continuous representations, such that $\overline{\rho} = \rho \pmod{\lambda} = \rho_0 \pmod{\lambda}$. Assume that ρ_0 is modular and that ρ is geometric. Assume further that the following properties hold.*

(1) *p is unramified in F.*
(2) $\mathrm{Im}\,\overline{\rho} \supseteq \mathrm{SL}_2(\mathbb{F}_p)$.
(3) *For all $v|p$, $\rho|_{G_{F_v}}$ and $\rho_0|_{G_{F_v}}$ are crystalline.*
(4) *For all $\sigma : F \hookrightarrow L$, $\mathrm{HT}_\sigma(\rho) = \mathrm{HT}_\sigma(\rho_0)$ contains two distinct elements differ by at most $p-2$.*

Then ρ is modular.

Recall that we have proved another modularity result in [92], which (partially) confirms the Shimura-Taniyama conjecture on the modularity of elliptic curves over $\mathbb{Q}$. Theorem 1.1 here is clearly stronger than the one in [92], as we have relaxed certain conditions and F could be totally real. We have explained the history of the Shimura-Taniyama conjecture and its relation with the Fermat's Last Theorem in [92]. In the last section of [92], we briefly discussed the limitation of the Taylor-Wiles-Diamond argument and some of the generalizations. These notes are the result of detailed studies on one of those generalizations.

In the second part of these notes [93], we studied Galois representation, especially the theory of (framed) deformations in detail. Using framed deformations, as opposed to merely deformations, we could deal with reducible residual Galois representations, and this is a huge improvement of [92]. We will use the results in [93] frequently in the following.

The contents of this article is as follows. In Section 2, we study the transition from modular forms to automorphic forms and the automorphic representations for general reductive groups. In Sections 3.1 and 3.3, we study some basic results on Langlands correspondence, especially the base change procedure, which enables us to simplify our situations tremendously. In these sections, we only describe the general picture but give no proofs, as it would take us too far away from our main goal. Nevertheless, in Sections 3.2 and 3.4, we study the special case, where the reductive group is associated with a quaternion algebra over a totally real field F, and provide complete proofs for most of the results.

In Section 4, we recall the Taylor-Wiles-Kisin construction from [50]. For completeness and for better understanding of results in Section 6, we also give a detail proof of the construction.

Section 5 is the key part of this article, where we construct the Taylor-Wiles-Kisin system explicitly and apply Kisin's result to obtain a proof of Theorem 1.1. For this we need basically all the preparations we have done and it also exhibits the power of the patching argument.

There are further things we could do with patching. In Section 6, we study the so called ultrapatching and make Proposition 4.1 into a "program" via pure algebraic method. The applications of ultrapatching is on the top of our to-do list.

The list of references is rather long in this article, as we include those references from [92]. We use the notation in [93] in the following.

2. From modular forms to automorphic forms

Historically the theory of automorphic forms began with modular and cusp forms for the group $\mathrm{SL}_2(\mathbb{Z})$. To understand the Langlands correspondence and the Jacquet-Langlands correspondence later in the notes, we review some general results on automorphic forms and automorphic representations. Here we will not be precise and will not provide proofs, details could be found in [5, 25, 53, 73].

2.1. **Modular forms for** $\mathrm{SL}_2(\mathbb{R})$**.**

The group $\mathrm{GL}_2(\mathbb{R})$ acts on $\mathbb{C}-\mathbb{R}$ via $g(z)=\frac{az+b}{cz+d}$, where $g=\begin{pmatrix} a & b \\ c & d \end{pmatrix}\in \mathrm{GL}_2(\mathbb{R})$ and $z\in\mathbb{C}-\mathbb{R}$. Define

$$j(g,z)=(cz+d)(\det g)^{1/2}.$$

In the rest of this section $j(g,z)$ will occur only in the form $j(g,z)^2$ as we only consider modular forms of level $\mathrm{SL}_2(\mathbb{Z})$ and of even weight.

Definition 2.1. Let $\mathbb{H}^+$ be the upper half plane. A *modular form* of weight k (an even integer) for $\mathrm{SL}_2(\mathbb{Z})$ is an analytic function $f:\mathbb{H}^+\to\mathbb{C}$ such that

(1) $f(\gamma(z))=j(\gamma,z)^k f(z)$ for all $f\in\mathrm{SL}_2(\mathbb{Z})$;
(2) f is analytic at ∞ in the following sense: since f is analytic and periodic under $z\mapsto z+1$, f has an expansion $f(z)=\sum_{n=-\infty}^{+\infty}c_n e^{2\pi i n z}$, then the analyticity condition at ∞ is that $c_n=0$ for $n<0$.

A *cusp form* is a modular form that vanishes at ∞ in the sense that $c_0=0$.

Remark 2.2. The analyticity condition at ∞ can be reformulated as the *slow-growth condition*

$$|f(x+iy)|\le Cy^N \text{ for some } C \text{ and } N \text{ as } y\to+\infty.$$

The additional condition that a modular form is a cusp form can be reformulated as the vanishing of an integral:

$$\int_0^1 f(x+iy)\,\mathrm{d}\,x = 0 \text{ for some or equivalently every } y > 0.$$

A cusp form f satisfies the *rapid-decrease condition* that for each N

$$|f(x+iy)| \leq Cy^{-N} \text{ for some } C \text{ as } y \to +\infty.$$

A cusp form of weight k can be expanded as $f(z) = \sum_{n=1}^{+\infty} c_n e^{2\pi i n z}$. The *L-function associated with* f is defined by

$$L(s,f) = \sum_{n=1}^{\infty} \frac{c_n}{n^s},$$

and it satisfies a functional equation relating the values at s and $k-s$. Let V_k be the space of cusp forms of weight k. Then V_k is finite dimensional and Hecke introduced what we now call *Hecke operators* on V_k. The Hecke operators commute and are simultaneously diagonalizable. The eigenfunctions all have $c_1 \neq 0$ and if c_1 is normalized to be 1 for an eigenfunction, then the corresponding L-function has an Euler product expansion and the product being taken over all primes.

Gelfand and Fomin were the first to notice that cusp forms could be realized as smooth vectors in representations of a certain ambient Lie group. We lift modular forms and cusp forms to $\mathrm{SL}_2(\mathbb{Z})\backslash \mathrm{SL}_2(\mathbb{R})$ and to $\mathrm{GL}_2(\mathbb{Z})\backslash \mathrm{GL}_2(\mathbb{R})$. See [5] for more details.

Given a modular form f as above, define

$$\phi_{f,\infty}(g) = f(g(i))j(g,i)^{-k} \tag{2.1}$$

for $g \in \mathrm{SL}_2(\mathbb{R})$. Then $\phi_{f,\infty}$ has the following properties:

(1) $\phi_\infty(\gamma g) = \phi_\infty(g)$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$;

(2) $\phi_\infty(gr(\theta)) = e^{-ik\theta}\phi_\infty(g)$ for all $r(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$;

(3) $\phi_\infty(g)$ satisfies the slow-growth condition that

$$|\phi_\infty(\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}\begin{pmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{pmatrix} r(\theta))| \leq Cy^N \text{ for some } C \text{ and } N \text{ as } y \to +\infty;$$

(4) $\Delta\phi_\infty = -\frac{k}{2}(\frac{k}{2}-1)\phi_\infty$ for a suitable normalization of the Casimir operator Δ of $\mathrm{SL}_2(\mathbb{R})$;

(5) if f is a cusp form, then ϕ_∞ is *cuspidal* in the sense that

$$\int_0^1 \phi_\infty(\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} g)\,\mathrm{d}\,x = 0 \text{ for all } g.$$

For the lifting to $\mathrm{GL}_2(\mathbb{Z})\backslash \mathrm{GL}_2(\mathbb{R})$, we start by extending f to $\mathbb{C}-\mathbb{R}$ by setting $f(-z) = f(z)$. Then we define $\phi_{f,\infty}(g)$ by equation (2.1) for $g \in \mathrm{GL}_2(\mathbb{R})$. The invariance property in (1) extends to be valid for all $\gamma \in \mathrm{GL}_2(\mathbb{Z})$, properties (2)-(5) are unchanged, and there is one new property:

- $\phi_\infty(zg) = \phi_\infty(g)$ for all z in the center of $\mathrm{GL}_2(\mathbb{R})$.

Remark 2.3. There are some other classical theories of automorphic forms that can be lifted to Lie groups in the same way. For example

- the theory of Maass forms [56] concerns certain non-holomorphic functions on $\mathbb{H}^+$, and these lift to $\mathrm{GL}_2(\mathbb{Z})\backslash \mathrm{GL}_2(\mathbb{R})$;
- a theory [25] begun by Hecke for modular form with level $\Gamma_0(N)$;
- the theory of Hilbert modular forms [34] leads to quotients of products of $\mathrm{GL}_2(\mathbb{R})$;
- the theory of Siegel modular forms [76] leads to quotients of real symplectic groups.

In each case the theory can be reinterpreted in an adelic setting, in which case one could see more symmetries in the space of forms. In the following, we explain the $\mathrm{SL}_2(\mathbb{Z})$ case with more detail.

Let $\mathbb{A}_\mathbb{Q}$ denote the ring of adeles of $\mathbb{Q}$. We have the following decomposition

$$\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}) = \mathrm{GL}_2(\mathbb{Q})\,\mathrm{GL}_2(\mathbb{R}) \prod_p \mathrm{GL}_2(\mathbb{Z}_p)$$

and each $g \in \mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ could be written as $g = \gamma g_\infty k_1$. Given a modular form f, define

$$\phi_f(g) = f(g_\infty(i)) j(g_\infty, i)^{-k} \text{ for } g \in \mathrm{GL}_2(\mathbb{A}_\mathbb{Q}).$$

The function $\phi := \phi_f$ on $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ has the properties that

(1) $\phi(\gamma g) = \phi(g)$ for all $\gamma \in \mathrm{GL}_2(\mathbb{Q})$;
(2) $\phi(gk_1) = \phi(g)$ for all $k_1 \in \prod_p \mathrm{GL}_2(\mathbb{Z}_p)$;
(3) $\phi(gr_\infty(\theta)) = e^{-ik\theta}\phi(g)$ for all $r_\infty(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ at the infinite place;
(4) as a function of the variable in the infinite place, ϕ satisfies the equation $\Delta\phi_\infty = -\frac{k}{2}(\frac{k}{2}-1)\phi_\infty$ for a suitable normalization of the Casimir operator Δ of $\mathrm{SL}_2(\mathbb{R})$;
(5) $\phi_\infty(zg) = \phi_\infty(g)$ for all scalar z in $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$
(6) $\phi_\infty(g)$ satisfies the slow-growth condition: for each $c > 0$ and compact subset ω of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$, there exist constants C and N such that

$$\phi(\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} g) \leq C|a|^N$$

for all $g \in \omega$ and $a \in \mathbb{A}_\mathbb{Q}^\times$ with $|a|_{\mathbb{A}_\mathbb{Q}} > c$;

(7) if f is a cusp form, then ϕ is *cuspidal* in the sense that

$$\int_{\mathbb{Q}\backslash\mathbb{A}_\mathbb{Q}} \phi(\begin{pmatrix}1 & x\\ 0 & 1\end{pmatrix} g)\,\mathrm{d}\,x = 0 \text{ for all } g \in \mathrm{GL}_2(\mathbb{A}_\mathbb{Q}).$$

The group $\mathrm{SL}_2(\mathbb{Z})$, relative to which f satisfies an invariance property, is captured by the compact group in property (2). The relevant identity is

$$\mathrm{GL}_2(\mathbb{Z}) = \mathrm{GL}_2(\mathbb{Q}) \cap (\mathrm{GL}_2(\mathbb{R}) \times \prod_p \mathrm{GL}_2(\mathbb{Z}_p)).$$

For the general congruence subgroup $\Gamma_0(N)$, the corresponding compact group that appears in property (2) is $\prod_p K'_p$, where

$$K'_p = \{\begin{pmatrix}a & b\\ c & d\end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z}_p) \mid v_p(c) \geq v_p(N)\}.$$

One sees that K'_p coincides with $\mathrm{GL}_2(\mathbb{Z}_p)$ for all p prime to N, and the relevant identity is

$$\Gamma_0(N) \cup \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \Gamma_0(N) = \mathrm{GL}_2(\mathbb{Q}) \cap (\mathrm{GL}_2(\mathbb{R}) \times \prod_p K'_p.$$

2.2. **Automorphic forms for general reductive group G.** The adelic setting is what one could generalize to arbitrary reductive groups (cf. [8]). Let F be a number field, let $\mathcal{O}$ be the ring of integers of F, let $\mathbb{A} = \mathbb{A}_F$ (resp. $\mathbb{A}^\infty$) be the ring of adeles of F (resp. finite adeles of F), let G be a reductive group over F such that $G(\mathbb{C})$ is connected. Let Z be a maximal F-split torus of the center of G. Let $G_\infty := G(F_\infty)$ be the archimedean component of $G(\mathbb{A})$, so that

$$G(\mathbb{A}) = G_\infty \times G(\mathbb{A}^\infty). \tag{2.2}$$

Let

$$\begin{aligned} K_\infty &= \text{a maximal compact subgroup of the Lie group } G_\infty\\ \mathfrak{g} &= \text{complexification of the (real) Lie algebra of } G_\infty\\ U(\mathfrak{g}) &= \text{universal enveloping algebra of } \mathfrak{g}\\ Z(\mathfrak{g}) &= \text{center of } \mathfrak{g}. \end{aligned}$$

Let K_1 be the open compact subgroup $G(\prod_{v \text{ finite}} \mathcal{O}_v)$ of $G(\mathbb{A}^\infty)$.

Definition 2.4. A function $f : G(\mathbb{A}) \to \mathbb{C}$ is *smooth* if it is continuous and, when viewed as a function of two variables (x, y) as in (2.2) ($x \in G_\infty$ and $y \in G(\mathbb{A}^\infty)$), it is smooth in x for each fixed y and is locally constant with compact support in y for each fixed x.

Let ρ be a finite-dimensional representation of K_∞, $J \subset Z(\mathfrak{g})$ be an ideal of finite codimension, and K be an open subgroup of K_1. A smooth function f on $G(\mathbb{A})$ is *automorphic* relative to (ρ, J, K) if

(1) $f(\gamma g) = f(g)$ for all $\gamma \in G(F)$;
(2) $f(gk) = f(g)$ for all $k \in K$;
(3) the span of the right translations of f by K_∞ is finite-dimensional, and every irreducible constituent of this representation of K_∞ is a constituent of ρ;
(4) $Jf = 0$, where J acts on f in the G_∞ variable;
(5) for each $y \in G(\mathbb{A}^\infty)$, the function $x \mapsto f(xy)$ on G_∞ satisfies a certain slow-growth condition. See [8] for the precise statement.

The set of automorphic functions relative to (ρ, J, K) will be denoted by $\mathcal{A}(\rho, J, K)$.

Remark 2.5. By properties (1) and (2), we may consider f as a function on the double quotient $G(F)\backslash G(\mathbb{A})/K$. Note that this double quotient has a structure of manifold and we have an identification

$$G(F)\backslash G(\mathbb{A})/K = \coprod_{c\in C} (\Gamma_c \backslash G_\infty). \tag{2.3}$$

Here C is a finite subset of $G(\mathbb{A})$ such that

$$G(\mathbb{A}) = \coprod_{c\in C} G(F)cG_\infty K,$$

and $\Gamma_c = G_\infty c K c^{-1} \cap G(F)$. The finiteness of C follows from our assumption that G is reductive, hence in particular G has the *strong approximation property*. The case $G = \mathrm{SL}_2/\mathbb{Q}$ and $K = K_1$ before is a simple example of (2.3) where $C = \{1\}$. The right side of (2.3) is more concrete than the left side, but part of the action is lost in working with the right side rather than with the adeles, e.g., the symmetry of $G(F_v)$ is missing.

Remark 2.6. When $G = \mathrm{GL}_1$, any grossencharacter is an example of an automorphic form relative to a suitable triple.

The following fundamental result is due to Harish-Chandra ([43, 44]), where it is proved in the setting of the right side of (2.3). The translation into the following form is in [8, Page 195].

Theorem 2.7. *For every triple (ρ, J, K), $\mathcal{A}(\rho, J, K)$ is finite-dimensional.*

Moreover, automorphic functions are closely related to L^p-functions. We have the following result [8, Page 191-195].

Theorem 2.8. *Let f be a function on $G(\mathbb{A})$ satisfying conditions (1)-(4) and*

$$f(zx) = \chi(z)f(x) \text{ for all } z \in Z(\mathbb{A}) \text{ and } x \in G(\mathbb{A}) \tag{2.4}$$

for some (unitary) character of $Z(F)\backslash Z(\mathbb{A})$, so that $|f|$ may be regarded as a function on $(Z(\mathbb{A})G(F))\backslash G(\mathbb{A})$. If $|f|$ is in $L^p((Z(\mathbb{A})G(F))\backslash G(\mathbb{A}))$ for some $p \geq 1$, then f satisfies condition (5) and hence is an automorphic form.

Remark 2.9 (Some geometry on the quotient space)*.* With G defined over F, let $X^*(G)_F$ be the set of all F-rational homomorphisms of G into GL_1. If $\chi \in X^*(G)_F$, then χ extends at each place to a continuous homomorphism $\chi_v : G(F_v) \to F_v^\times$. Let $\chi_\mathbb{A} : G(\mathbb{A}) \to \mathbb{A}^\times$ be the product of χ_v. Then $|\chi_\mathbb{A}|_\mathbb{A}$ is a homomorphism of $G(\mathbb{A})$ into $\mathbb{R}_+^\times$. Define $G(\mathbb{A})^1 := \bigcap_{\chi \in X^*(G)_F} \mathrm{Ker}\, |\chi_\mathbb{A}|_\mathbb{A}$.

Recall that $G(\mathbb{C})$ is connected in our setting. In this case, the group $G(F)$ lies in $G(\mathbb{A})^1$ and the quotient $G(F)\backslash G(\mathbb{A})^1$ has finite volume. Moreover, $G(F)\backslash G(\mathbb{A})^1$ is compact if and only if every unipotent element of $G(F)$ belongs to the radical of $G(F)$. See for example [3, 4].

An example of a nonabelian G for which compactness of $G(F)\backslash G(\mathbb{A})^1$ follows from this result is the multiplicative group $D^\times$ of a finite-dimensional division algebra of F with center F. This is exactly the case we will study later.

Definition 2.10. A *cusp form* is an automorphic form f such that (2.4) holds for some unitary character χ of $Z(F)\backslash Z(\mathbb{A})$ and such that

$$\int_{N(F)\backslash N(\mathbb{A})} f(ng)\, \mathrm{d}\, n = 0 \tag{2.5}$$

for the unipotent radical N of every parabolic subgroup of G and for all $g \in G(\mathbb{A})$. Let $\mathcal{A}^0(\rho, J, K)$ denote the space of cusp forms relative to (ρ, J, K).

Remark 2.11*.* For $G = \mathrm{GL}_1$, the condition (2.5) is empty, and therefore all unitary grossencharacters are cusp forms for GL_1.

The classical analytic cusp forms relative to $\mathrm{SL}_2(\mathbb{Z})$ yield cusp forms for $G = \mathrm{GL}_2/\mathbb{Q}$ in the sense of Definition 2.10. Similar relation holds for congruence groups $\Gamma_0(N)$.

Theorem 2.12. *Let a smooth function f on $G(\mathbb{A})$ satisfy (1)-(4) above, as well as the cuspidal condition* (2.5) *and the condition* (2.4) *for some (unitary) character of $Z(F)\backslash Z(\mathbb{A})$. Then the following conditions are equivalent:*

1. *f satisfies (5) and hence is a cusp form;*
2. *f is bounded;*
3. *$|f|$ is in $L^2((Z(\mathbb{A})G(F))\backslash G(\mathbb{A}))$.*

2.3. **Automorphic representations.** We want to define notion of an automorphic representation of $G(\mathbb{A})$. Put $\mathcal{A} = \cup \mathcal{A}(\rho, J, K)$. The idea is that an automorphic representation is any irreducible subquotient of $\mathcal{A}$, but the trouble is that $\mathcal{A}$ need not be mapped to itself under right translation of $G(\mathbb{A})$. Specifically, right translation by $G(\mathbb{A})$ does not preserve the K_∞-finiteness in general. The idea is to make $\mathcal{A}$ into a module for an algebra $\mathcal{H}$ (the *Hecke algebra*) that reflects the action by $G(F_v)$ for each finite place v and reflects the action by $U(\mathfrak{g})$ and K_∞ at the infinite place. The following is a summary of Knapp [53]. See [28] for the detail construction, see [2] for the relation between representations of real reductive Lie groups and $(\mathfrak{g}, K)$-modules.

For each finite place v, let $\mathcal{H}_v$ be the space of all complex-valued locally constant functions of compact support on $G(F_v)$. Let $f, f' \in \mathcal{H}_v$, define

$$(f * f')(x) = \int_{G(F_v)} f(y) f'(y^{-1}x) \,\mathrm{d}\, y = \int_{G(F_v)} f(y^{-1}) f'(yx) \Delta(y^{-1}) \,\mathrm{d}\, y,$$

and

$$f^*(x) = \Delta(x^{-1}) \overline{f(x^{-1})}.$$

These two operations make $\mathcal{H}_v$ into a Banach $*$-algebra with no unity. Haar measure on $G(F_v)$ is to be normalized so that $G(\mathcal{O}_v)$ has measure 1. Then the characteristic function I_v of $G(\mathcal{O}_v)$ is an idempotent in $\mathcal{H}_v$. Easy computation shows that the normalized characteristic function of each open subgroup of $G(\mathcal{O}_v)$ is an idempotent and they form a directed system of idempotents.

Definition 2.13. An $\mathcal{H}_v$-module is *approximately unital* if, for each member of the module, all idempotents corresponding to sufficiently small open subgroups of $G(\mathcal{O}_v)$ act as the identity.

A $G(F_v)$-representation is *smooth* if each member of the representation space is fixed by some open compact subgroup of $G(F_v)$.

It is easy to see that smooth $G(F_v)$-representations correspond to approximately unital $\mathcal{H}_v$-modules. Such a representation is called *addmissible* if the set of vectors fixed by any open compact subgroup is finite-dimensional.

There is a natural way of forming a restricted tensor product of the algebras $\mathcal{H}_v$ with respect to the idempotents I_v. The resulting algebra $\mathcal{H}_f$ is the part of $\mathcal{H}$ corresponding to the finite places of F and is generated by product functions that equal I_v at almost every place. A tuple of local idempotents, one for each $\mathcal{H}_v$ with almost all of them being I_v, yields another idempotent in $\mathcal{H}_f$, and the idempotents obtained in this way are indexed by a directed set.

Definition 2.14. A right $\mathcal{H}_f$-module is *smooth* if each member of the module is fixed by all idempotents corresponding to members of the directed set that are sufficiently large. The module is *admissible* if the set of vectors fixed by any of these idempotents is finite-dimensional.

Next let $\mathcal{H}_\infty$ be the convolution algebra of all K_∞-finite distributions on G_∞ that are supported on K_∞. It contains a directed family of idempotents as follows, constructed via the Peter-Weyl Theorem. Let $\mathrm{d}\,k$ denote the normalized Haar measure on K_∞. For each class of irreducible representations τ of K_∞, let χ_τ be the character and let d_τ be the degree. The directed family of idempotents is indexed by all finite subsets of τ's, the idempotent corresponding to a given set being the sum of $d_\tau\chi_\tau$ for all τ in the set.

Definition 2.15. A right $\mathcal{H}_\infty$-module is *approximately unital* if, for each member of the module, all sufficiently large idempotents act as the identity. Such a module is *admissible* if the set of vectors fixed by any of these idempotents is finite-dimensional, i.e., if each K_∞-type has finite multiplicity.

Remark 2.16. As explained in [54], $(\mathfrak{g}, K_\infty)$-modules correspond exactly to approximately unital $\mathcal{H}_\infty$-modules.

Define $\mathcal{H} = \mathcal{H}_\infty \otimes \mathcal{H}_f$. Smoothness and admissibility of right $\mathcal{H}$-modules are defined using idempotents that are pure tensors from $\mathcal{H}_\infty$ and $\mathcal{H}_f$. Then $\mathcal{A}$ is a sooth right $\mathcal{H}$-module. An *automorphic representation* of $\mathcal{H}$ is any irreducible subquotient of $\mathcal{A}$. Similarly, if we put $\mathcal{A}^0 = \cup\mathcal{A}^0(\rho, J, K)$, then a *cuspidal automorphic representation* of $\mathcal{H}$ is any irreducible subquotient of $\mathcal{A}^0$.

If f is an automorphic form, then by Theorem 2.7, $f * \mathcal{H}$ is a smooth admissible $\mathcal{H}$-module. It follows that every automorphic representation of $\mathcal{H}$ is smooth and admissible. Such representations are commonly called automorphic representations of $G(\mathbb{A})$ even though not all of $G(\mathbb{A})$ really acts.

Definition 2.17. A topologically irreducible $G(\mathbb{A})$-module is called *automorphic* if its underlying space of smooth vectors is an automorphic representation of $\mathcal{H}$.

According to [28, Theorem 4], if χ is any (unitary) character of $Z(F)\backslash Z(\mathbb{A})$, then any $G(\mathbb{A})$-invariant irreducible closed subspace of

$$\begin{aligned}&L^2(G(F)\backslash G(\mathbb{A}))_\chi\\ &\quad = \{f \mid f \in L^2(G(F)\backslash G(\mathbb{A})) \text{ and } f(zx) = \chi(z)f(x) \text{ for } z \in Z(\mathbb{A}),\ x \in G(\mathbb{A})\}\end{aligned}$$

is automorphic in this sense. The following result is due to Gelfand and Piatetski-Shapiro [38].

Theorem 2.18. *The subspace of cuspidal functions in $L^2(G(F)\backslash G(\mathbb{A}))_\chi$ decomposes discretely with finite multiplicity. Consequently whenever f is a cusp form, $f * \mathcal{H}$ is a finite direct sum of cuspidal automorphic representations.*

Remark 2.19. It follows from the theorem that cuspidal automorphic representations are unitarizable. That is, they are the underlying smooth representations for irreducible unitary representations of $G(\mathbb{A})$. One principal object of Langlands correspondence is the set of equivalence classes of smooth admissible representations of $G(\mathbb{A})$. Being unitarizable is an important property of the cuspidal automorphic representations.

Remark 2.20. Classical cusp forms for $\mathrm{GL}_2/\mathbb{Q}$ lead to cusp forms in the adelic setting by Theorem 2.12. Those whose L-function has Euler product expansions (i.e. eigenforms) lead to adelic cusp forms that generate single irreducible cuspidal automorphic representations.

3. Automorphic forms on quaternion algebras

In this section, following the general idea in last section, we study automorphic forms on quaternion algebras in detail.

3.1. **Representations of** $\mathrm{GL}_2(E)$**.** In this section, E is a local field with mixed characteristic with valuation ring $\mathcal{O}$. We fix an uniformizer ϖ_E of E. We briefly review the representations of $\mathrm{GL}_2(E)$ (cf. [5, 14]). For simplicity, let $G = \mathrm{GL}_2(E)$ and $K = \mathrm{GL}_2(\mathcal{O})$, so that K is a maximal open compact subgroup of G. Let $B = NM = MN$, where

$$M = \{m = m(a) = \begin{pmatrix} a_1 & \\ & a_2 \end{pmatrix}\} \text{ and } N = \{n = n(x) = \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix}\},$$

where $a_1,\ a_2 \in E^\times$ and $x \in E$. Let Z be the center of G.

Definition 3.1. Let (π, V) be a representation of G. We say that π is *smooth* if for any $v \in V$, the stabilizer of v in G is open. We say that a smooth π is *admissible* if for any open compact subgroup $U \subset G$, V^U is finite-dimensional.

An irreducible admissible representation (π, V) of G is called *unramified* if $V^K \neq 0$.

For each pair $\mu = (\mu_1, \mu_2)$ of characters of $E^\times$, there is an induced representation

$$I(\mu) = \{f : G \to \mathbb{C} \mid f(bg) = \mu(b)\delta(b)^{1/2} f(g)\},$$

where f is smooth (i.e. right invariant under an open subgroup $K' \subset G$), and for $b = nm(a)$,

$$\mu(b) = \mu_1(a_1)\mu_2(a_2) \text{ and } \delta(b) = |a_1/a_2|.$$

Note that δ is the modulus for the adjoint action of B on N, i.e.,

$$\mathrm{d}(bnb^{-1}) = \delta(b)\,\mathrm{d}\,n$$

for a Haar measure $\mathrm{d}\,n$ on N. Thus the induction is normalized so that a pair (μ_1, μ_2) of unitary characters yields a unitarizable representation $I(\mu)$. As we have the Iwasawa decomposition $G = BK$, the functions in $I(\mu)$ are determined by their restriction to K and we have

$$I(\mu)|_K \cong I_{K\cap B}^K(\mu).$$

It is also easy to see that the central character of $I(\mu)$ is $\mu_1\mu_2$. The following two theorems provide a partial classification of admissible $\mathrm{GL}_2(E)$-representations.

Theorem 3.2. *With notation as above, the following claims hold.*

(1) $I(\mu)$ *is irreducible if and only if* $\mu_1\mu_2^{-1} \neq \omega_{\pm 1}$, *where* $\omega_s(x) = |x|^s$.
(2) *If* $\mu_1\mu_2^{-1} = \omega_1$, *then* $I(\mu)$ *has a one-dimensional quotient on which* G *acts by the character* $\chi \circ \det$, *where* $\mu = (\chi\omega_{1/2}, \chi\omega_{-1/2})$, *and an infinite dimensional irreducible subrepresentation* $\sigma(\mu)$.
(3) *If* $\mu_1\mu_2^{-1} = \omega_{-1}$, *then* $I(\mu)$ *has a one-dimensional subrepresentation on which* G *acts by the character* $\chi \circ \det$, *where* $\mu = (\chi\omega_{-1/2}, \chi\omega_{1/2})$, *and an infinite dimensional irreducible quotient* $\sigma(\mu)$.
(4) *The only equivalences among these representations are the following. Let* $\mu' = (\mu_2, \mu_1)$.
   - *If* $\mu_1\mu_2^{-1} \neq \omega_{\pm 1}$, *then* $I(\mu) \cong I(\mu')$.
   - *If* $\mu_1\mu_2^{-1} = \omega_{\pm 1}$, *then* $\sigma(\mu) \cong \sigma(\mu')$.

Theorem 3.3. *For an irreducible admissible representation* (π, V) *of* $G = \mathrm{GL}_2(E)$, *the following are equivalent.*

(1) $\mathrm{Hom}_G(\pi, I(\mu)) = 0$ *for all* μ.
(2) *The matrix coefficients of* π *are compactly supported modulo* Z.

The representations satisfying the conditions in Theorem 3.3 are the *supercuspidal* representations of G. Together with the *irreducible principal series* representations $I(\mu)$, the *special* representations $\sigma(\mu)$ and the *one-dimensional* representations $\chi \circ \det$, they give all of the irreducible admissible representations of G up to isomorphism. The special representations and the supercuspidal representations are called *discrete series*.

If the characters μ_1 and μ_2 are unramified, so that

$$\mu_j(x) = t_j^{\mathrm{ord}(x)}$$

for some $(t_1, t_2) \in (\mathbb{C}^\times)^2$, then $\mu|_{(B\cap K)} = 1$ and we have an K-isomorphism

$$I(\mu)|_K \cong C^\infty(B \cap K\backslash K). \tag{3.1}$$

Therefore, we have $\dim_{\mathbb{C}} I(\mu)^K = 1$. In fact, this construction accounts for all unramified representations.

Theorem 3.4. *The following claims hold.*

1. *For every pair $(t_1, t_2) \in (\mathbb{C}^\times)^2$, there is an irreducible admissible unramified representation $\pi(t_1, t_2)$ of G.*
2. *Every irreducible admissible unramified representation of G is isomorphic to one $\pi(t_1, t_2)$. The only equivalence between such representations is*
$$\pi(t_1, t_2) \cong \pi(t_2, t_1).$$
3. *The special representations and the supercuspidal representations are ramified.*

The parametrization of the unramified representations of G can be expressed as follows, which is a special case of the so called *Satake parametrization* for an arbitrary connected reductive group G (cf. [41]).

Corollary 3.5. *There is a bijection between isomorphism classes of irreducible admissible unramified representations of G and semisimple conjugacy classes in complex group $G^\vee := \mathrm{GL}_2(\mathbb{C})$, given by*

$$\pi = \pi(t_1, t_2) \leftrightarrow t(\pi) := \textit{conjugacy class of } \begin{pmatrix} t_1 & \\ & t_2 \end{pmatrix}.$$

The local Langlands correspondence provides a unique family of bijections rec_E from the set of irreducible admissible representations of $\mathrm{GL}_n(E)$ to the set of n-dimensional Frobenius semisimple Weil-Deligne representations of W_E over $\mathbb{C}$, satisfying a list of properties. In order to be uniquely determined, one needs to formulate the correspondence for all n at once, and the properties are expressed in terms of L- and ϵ-factors, neither of which we have defined and we will need later. We state the properties of the correspondence that we need to use.

Theorem 3.6. *The rec_E for $n = 1$ is given by $\mathrm{rec}_E(\pi) = \pi \circ \mathrm{Art}_E^{-1}$. Here $\mathrm{Art} : G_E \to E^\times$ is the Artin reciprocity map.*

The rec_E for $n = 2$ satisfies the following conditions.

1. *If χ is a smooth character, then $\mathrm{rec}_E(\pi \otimes (\chi \circ \det)) = \mathrm{rec}_E(\pi) \otimes \mathrm{rec}_E(\chi)$.*
2. *If $\mu_1\mu_2^{-1} \neq \omega_{\pm 1}$, then $\mathrm{rec}_E(I(\mu)) = \mu_1 \oplus \mu_2$.*

(3) *If* $\mu_1\mu_2^{-1} = \omega_{\pm 1}$, *i.e.,* $\mu = (\chi\omega_{1/2}, \chi\omega_{-1/2})$ *or* $\mu = (\chi\omega_{-1/2}, \chi\omega_{1/2})$, *then* $\mathrm{rec}_E(\sigma(\mu)) = \chi \oplus \chi\omega_1$ *with* $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. *This is the only case with* N *nontrivial. In this case,* $\mathrm{rec}_E(\sigma(\mu))$ *is indecomposable.*

(4) *If* $\pi = \chi \circ \det$, *then* $\mathrm{rec}_E(\pi) = \chi\omega_{1/2} \oplus \chi\omega_{-1/2}$.

(5) π *is discrete series if and only if* $\mathrm{rec}_K(\pi)$ *is indecomposable, is cuspidal if and only if* $\mathrm{rec}_E(\pi)$ *is irreducible.*

Remark 3.7 (Hecke operators). Let ψ be a compactly supported $\mathbb{C}$-valued function on $\mathrm{GL}_2(\mathcal{O})\backslash \mathrm{GL}_2(E)/ \mathrm{GL}_2(\mathcal{O})$. Concretely, these are functions which vanish outside of a finite number of double cosets $\mathrm{GL}_2(\mathcal{O})g\,\mathrm{GL}_2(\mathcal{O})$. The set of such functions is in fact a ring, with the multiplication being given by convolution (cf. Section 2.3, where we consider the non-unital Hecke algebra). To be precise, we fix μ the (left and right) Haar measure on $\mathrm{GL}_2(E)$ such that $\mu(\mathrm{GL}_2(\mathcal{O})) = 1$ and define

$$(\psi_1 * \psi_2)(x) = \int_{\mathrm{GL}_2(E)} \psi_1(g)\psi_2(g^{-1}x)\,\mathrm{d}\,\mu_g.$$

From our assumption on ψ, this integral is just a finite sum. Denote this ring by $\mathcal{H}_E$. Then in this case, it is isomorphic to $\mathbb{C}[T, S^{\pm 1}]$, where T is the characteristic function of the double coset

$$\mathrm{GL}_2(\mathcal{O}) \begin{pmatrix} \varpi_E & 0 \\ 0 & 1 \end{pmatrix} \mathrm{GL}_2(\mathcal{O})$$

and S is the characteristic function of the double coset

$$\mathrm{GL}_2(\mathcal{O}) \begin{pmatrix} \varpi_E & 0 \\ 0 & \varpi_E \end{pmatrix} \mathrm{GL}_2(\mathcal{O}).$$

The algebra $\mathcal{H}_E$ acts on an irreducible admissible $\mathrm{GL}_2(E)$-representation π. Given $\psi \in \mathcal{H}_E$, we have a linear map $\pi(\psi)$ given by

$$\begin{aligned} \pi(\psi) : \pi &\to \pi^{\mathrm{GL}_2(\mathcal{O})} \subset \pi \\ w &\mapsto \int_{\mathrm{GL}_2(E)} \psi(g)\pi(g)w\,\mathrm{d}\,\mu_g. \end{aligned}$$

In particular, if π is an unramified irreducible admissible representation, then $\pi^{\mathrm{GL}_2(\mathcal{O})}$ is one dimensional and $\pi(\psi)$ acts via a scalar on this one-dimensional space. As we have a classification of π, we could compute this action explicitly.

First note that we have two decompositions

$$\mathrm{GL}_2(\mathcal{O}) \begin{pmatrix} \varpi_E & 0 \\ 0 & \varpi_E \end{pmatrix} \mathrm{GL}_2(\mathcal{O}) = \begin{pmatrix} \varpi_E & 0 \\ 0 & \varpi_E \end{pmatrix} \mathrm{GL}_2(\mathcal{O}) \tag{3.2}$$

and

$$\mathrm{GL}_2(\mathcal{O})\begin{pmatrix}\varpi_E & 0\\ 0 & 1\end{pmatrix}\mathrm{GL}_2(\mathcal{O}) = (\coprod_{a\in k_E}\begin{pmatrix}\varpi_E & \tilde{a}\\ 0 & 1\end{pmatrix}\mathrm{GL}_2(\mathcal{O}))\coprod\begin{pmatrix}1 & 0\\ 0 & \varpi_E\end{pmatrix}\mathrm{GL}_2(\mathcal{O}), \tag{3.3}$$

where $\tilde{a}\in\mathcal{O}$ is a lifting of a. Denote by q the cardinality of the residue field of E. Then it is easy to check that

(1) Suppose that $\pi = (\chi|\cdot|^{1/2})\circ\det$ with χ unramifed. Then $\pi^{\mathrm{GL}_2(\mathcal{O})} = \pi$, and S and T acts via $\chi(\varpi_E)^2q^{-1}$ and $\chi(\varpi_E)(q^{1/2}+q^{-1/2})$ respectively.
(2) Suppose that χ_1 and χ_2 are unramified characters and that $\chi_1\chi_2^{-1}\neq|\cdot|^{\pm1}$. Let $\pi = I(\mu_1,\mu_2)$. Then from equation (3.1), $\pi^{\mathrm{GL}_2(\mathcal{O})}$ is spanned by a function ϕ_0 with $\phi_0(\begin{pmatrix}a & b\\ 0 & d\end{pmatrix}) = \chi_1(a)\chi_2(d)|a/d|^{1/2}$. Moreover, S and T acts on $\pi^{\mathrm{GL}_2(\mathcal{O})}$ via $\chi_1\chi_2(\varpi_E)$ and $q^{1/2}(\chi_1(\varpi_E)+\chi_2(\varpi_E))$ respectively.

3.2. **Notation and definition.** In this section, we construct antomorphic forms on definite quaternion algebras (cf. [36, 84]). Let F be a totally real field. Let D be a quaternion algebra over F, i.e. central simple F-algebra of dimension 4. Denote by $S(D)$ the set of ramified places of D over F, i.e. those for which $D\otimes_F F_v$ is a division algebra, (equivalently, is not isomorphic to $M_2(F_v)$). Then $S(D)$ is a finite set of places of F with even cardinality and it determines D up to isomorphism. In particular, $S(D)=\emptyset$ if and only if $D = M_2(F)$.

Fix a maximal order $\mathcal{O}_D$ of D, that is, a $\mathbb{Z}$-subalgebra of D which is finitely generated as a $\mathbb{Z}$-module and for which $\mathcal{O}_D\otimes_{\mathbb{Z}}\mathbb{Q}\cong D$. For example, if $D = M_2(F)$, one may take $\mathcal{O}_D = M_2(\mathcal{O}_F)$. For $v\notin S(D)$ and finite, fix an isomorphism $D\otimes_F F_v\xrightarrow{\sim} M_2(F_v)$ such that it induces an isomorphism $\mathcal{O}_D\otimes_{\mathcal{O}_F}\mathcal{O}_{F_v}\xrightarrow{\sim} M_2(\mathcal{O}_{F_v})$.

Let $G := G_D$ be the algebraic group over $\mathbb{Q}$ such that for any $\mathbb{Q}$-algebra R, $G(R) = (D\otimes_{\mathbb{Q}} R)^\times$. For each infinite place $v|\infty$ of F, we define a subgroup U_v of $G(F_v) = (D\otimes_F F_v)^\times$ as follows.

- If $v\in S(D)$, then $G(F_v) = (D\otimes_F F_v)^\times = \mathbf{H}^\times$, where $\mathbf{H}$ is the Hamilton quaternion algebra. Define $U_v = (D\otimes_F F_v)^\times$.
- If $v\notin S(D)$, then $G(F_v) = (D\otimes_F F_v)^\times = \mathrm{GL}_2(\mathbb{R})$. Define $U_v = \mathbb{R}^\times\,\mathrm{SO}_2(\mathbb{R})$, the group generated by the center and the maximal compact subgroup.

If $\gamma = \begin{pmatrix}a & b\\ c & d\end{pmatrix}\in\mathrm{GL}_2(\mathbb{R})$ and $z\in\mathbb{C}-\mathbb{R}$, we let $j(\gamma,z) = cz+d$. Note that here j is a little different from the one we introduced in last section, but this is only for writing convenience. It is easy to check that if we have two matrices γ and δ, then

$$j(\gamma\delta, z) = j(\gamma, \delta z) j(\delta, z).$$

For each $v|\infty$, fix $k_v \geq 2$ and $\eta_v \in \mathbb{Z}$ such that $k_v + 2\eta_v - 1 = w$ is independent of v. These will be the weights of our modular forms. We define a representation (τ_v, W_v) of U_v over $\mathbb{C}$.

- If $v \in S(D)$, we have $U_v \hookrightarrow \mathrm{GL}_2(\overline{F_v}) \cong \mathrm{GL}_2(\mathbb{C})$ which acts on $\mathbb{C}^2$. We let (τ_v, W_v) be the representation

$$(\mathrm{Sym}^{k_v-2}\,\mathbb{C}^2) \otimes (\wedge^2\mathbb{C}^2)^{\eta_v}.$$

- If $v \notin S(D)$, we have $U_v \cong \mathbb{R}^\times\,\mathrm{SO}_2(\mathbb{R})$. We let (τ_v, W_v) be the one dimensional representation given by

$$\tau_v(\gamma) = j(\gamma, i)^{k_v} (\det\gamma)^{\eta_v - 1}.$$

We write $U_\infty = \prod_{v|\infty} U_v$, $W_\infty = \otimes_{v|\infty} W_v$, $\tau_\infty = \otimes_{v|\infty}\tau_v$. Let $\mathbb{A} = \mathbb{A}_\mathbb{Q}$ be the adeles of $\mathbb{Q}$ and $\mathbb{A}^\infty$ the finite adeles. Define $S_{D,k,\eta}$ to be the space of functions $f : G(\mathbb{Q})\backslash G(\mathbb{A}) \to W_\infty$ which satisfy the following conditions:

(1) $\phi(gu_\infty) = \tau_\infty(u_\infty)^{-1}\phi(g)$ for all $u_\infty \in U_\infty$ and $g \in G(\mathbb{A})$.
(2) There is a nonempty open subset $U^\infty \subset G(\mathbb{A}^\infty)$ such that $\phi(gu) = \phi(g)$ for all $u \in U^\infty$ and $g \in G(\mathbb{A})$.
(3) Let S_∞ denote the set of infinite places of F. If $g \in G(\mathbb{A}^\infty)$ then the function

$$\begin{aligned} h_\infty : (\mathbb{C} - \mathbb{R})^{S_\infty - S(D)} &\to W_\infty \\ (i,\ \ldots,\ i) &\mapsto \tau_\infty(h_\infty)\phi(gh_\infty) \end{aligned}$$

is holomorphic. Note that this function is well-defined by the first condition, as U_∞ is the stabilizer of $(i,\ \ldots,\ i)$.
(4) If $S(D) = \emptyset$, then for all $g \in G(\mathbb{A}) = \mathrm{GL}_2(\mathbb{A}_F)$, we have

$$\int_{F\backslash\mathbb{A}_F} \phi(\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} g)\,\mathrm{d}\,x = 0.$$

If in addition we have $F = \mathbb{Q}$, then we furthermore demand that for all $g \in G(\mathbb{A}^\infty)$, $h_\infty \in \mathrm{GL}_2(\mathbb{R})^+$, the function $\phi(gh_\infty)|\operatorname{Im}(h_\infty i)|^{k/2}$ is bounded on $\mathbb{C} - \mathbb{R}$.

There is a natural action of $G(\mathbb{A}^\infty)$ on $S_{D,k,\eta}$ by right translation, i.e., $(g\phi)(x) := \phi(xg)$. Note that here we do not consider the right translation of G_∞, the problem mentioned at the beginning of Section 2.3 dose not appear here.

Remark 3.8. If D is a definite quaternion algebra, i.e. $S_\infty \subset S(D)$, the situation is particularly simple, since the double quotient

$$\Gamma(\mathbb{Q})\backslash G(\mathbb{A})/G_\infty U$$

is a finite set, where $U \subset G(\mathbb{A}^\infty)$ is an open subgroup. When proving modularity lifting theorems for rank two Galois representations, we will be able to reduce to this simple case as we shall see in Section 3.3.9.

Definition 3.9. A representation (π, V) of $G(\mathbb{A}^\infty)$ is called *admissible* if

- for any $x \in V$, the stabilizer of x is open, and
- for any $U \subset G(\mathbb{A}^\infty)$ an open subgroup, $\dim_{\mathbb{C}} V^U < \infty$.

With this definition, $S_{D,k,\eta}$ is a semisimple admissible representation of $G(\mathbb{A}^\infty)$. As explained in [28], irreducible admissible representations of $G(\mathbb{A}^\infty)$ are restricted tensor products of irreducible admissible representations of $G(F_v)$ (v finite). Let us first recall the general construction. Let I be an indexing set and let V_i be a $\mathbb{C}$-vector space ($i \in I$). Suppose that we are given $0 \neq e_i \in V_i$ for almost all $i \in I$. Then we define the *restricted tensor product*

$$\otimes'_{e_i} V_i := \varinjlim_{J \in I}(\otimes_{i \in J} V_i),$$

where the colimit is over the finite subsets $J \subset I$ containing all the places for which e_i is not defined, the transition maps for the colimit are given by tensoring with the e_i. It is easy to check that $\otimes'_{e_i} V_i \cong \otimes'_{f_i} V_i$ if for almost all i, e_i and f_i span the same line.

Now if π_v is an irreducible smooth (so admissible) representation of $(D \otimes_F F_v)^\times$. Assume that $\pi_v^{\mathrm{GL}_2(\mathcal{O}_{F_v})} \neq \{0\}$ for all most all v. These are called *unramified* representations, and in this case, we must have $\dim_{\mathbb{C}} \pi_v^{\mathrm{GL}_2(\mathcal{O}_{F_v})} = 1$ (cf. Section 3.1). Then $\otimes' \pi_v$ is an irreducible admissible representation of $G(\mathbb{A}^\infty)$, where the restricted tensor product is taken with respect to the one dimensional spaces $\pi_v^{\mathrm{GL}_2(\mathcal{O}_{F_v})}$. Moreover any irreducible admissible representation of $G(\mathbb{A}^\infty)$ arises in this way for unique π_v.

Definition 3.10. The irreducible constituents of $S_{D,k,\eta}$ are called the *cuspidal automorphic representations* of $G(\mathbb{A}^\infty)$ of weight (k, η).

The Hecke algebra is easy to construct in this case (cf. Section 2.3). For each finite place v of F we choose $U_v \subset G(F_v)$ a open compact subgroup, such that $U_v = \mathrm{GL}_2(\mathcal{O}_{F_v})$ for almost all v. Let μ_v be a Haar measure on $G(F_v)$ normalized so that $\mu_v(G(\mathcal{O}_{F_v})) = 1$. Then there is a unique Haar measure μ on $G(\mathbb{A}^\infty)$ such that for any U_v as above, if we set $U = \prod_v U_v \subset G(\mathbb{A}^\infty)$, then $\mu(U) = \prod_v \mu_v(U_v)$. Then there is a decomposition

$$C_c(U\backslash G(\mathbb{A}^\infty)/U)\mu \cong \otimes'_{\mathbf{1}_{U_v}\mu_v} C_c(U_v\backslash G(F_v)/U_v)\mu_v,$$

which decomposes the (finite) global Hecke algebra as a restricted product of the local Hecke algebras. Certainly, the actions of these Hecke algebras are compatible with the decomposition $\pi = \otimes' \pi_v$.

3.3. **Langlands correspondence: some facts.** Now we review some important facts on automorphic representations and the Jacquet-Langlands correspondence, which enable us to reduce the proof of modularity results to a relatively simple situation.

3.3.1. *Multiplicity one for* GL_2. Suppose that $S(D) = \emptyset$. Then every irreducible constituent of $S_{D,k,\eta}$ has multiplicity one. Moreover, if π (resp. π') is a cuspidal automorphic representation of weight (k, η) (resp. (k', η')) such that $\pi_v \cong \pi'_v$ for almost all v, then $k = k'$, $\eta = \eta'$, and $\pi \cong \pi'$. (cf. [8, 5, 72, 46].)

3.3.2. *The theory of newforms.* Suppose that $S(D) = \emptyset$. If $\mathfrak{n}$ is an ideal of $\mathcal{O}_F$, write

$$U_1(\mathfrak{n}) = \{g \in \mathrm{GL}_2(\mathcal{O}_F) \mid g \equiv \begin{pmatrix} * & * \\ 0 & 1 \end{pmatrix} \pmod{\mathfrak{n}}\}.$$

If π is a cuspidal automorphic representation of $G(\mathbb{A}^\infty)$, then there is a unique ideal $\mathfrak{n}$ such that $\pi^{U_1(\mathfrak{n})}$ is one-dimensional, and $\pi^{U_1(\mathfrak{m})} \neq 0$ if and only if $\mathfrak{n}|\mathfrak{m}$. We call $\mathfrak{n}$ the *conductor* of π, $\pi^{U_1(\mathfrak{n})}$ the space of newforms. (cf. [8, 18].)

3.3.3. *The local Jacquet-Langlands correspondence.* Let E be a local field with mixed characteristic $(0, p)$. Analogous to the theory of admissible representations of $\mathrm{GL}_2(E)$, there is a theory of admissible representations of $D^\times$, where D/E is a nonsplit quaternion algebra. Since $D^\times/E^\times$ is compact, any irreducible admissible representation of $D^\times$ is finite dimensional. There is a bijection JL, the *local Jacquet-Langlands correspondence*, from the irreducible admissible representations of $D^\times$ to the discrete series representations of $\mathrm{GL}_2(E)$. (cf. [45, 8, 14].)

3.3.4. *The global Jacquet-Langlands correspondence.* Let D be a quaternion algebra over F such that $S(D)$ is nonempty. We have the following facts on cuspidal automorphic representations of $G(\mathbb{A}^\infty)$ (cf. [45, 8, 14]).

(1) The only finite dimensional cuspidal automorphic representation of $G(\mathbb{A}^\infty)$ are one-dimensional representations which factor through the reduced norm.
(2) There is a bijection JL from the infinite-dimensional cuspidal automorphic representaions of $G(\mathbb{A}^\infty)$ of weight (k, η) to the cuspidal automorphic representations of $\mathrm{GL}_2(\mathbb{A}^\infty)$ of weight (k, η) which are discrete series for all finite places $v \in S(D)$.
(3) The local and global Jacquet-Langlands correspondence are compatible in the following sense:

- if $v \notin S(D)$, then $\mathrm{JL}(\pi)_v = \pi_v$;
- if $v \in S(D)$, then $\mathrm{JL}(\pi)_v = \mathrm{JL}(\pi_v)$.

3.3.5. *Galois representations associated with automorphic representations.* Let π be a regular algebraic cuspidal automorphic representation of $\mathrm{GL}_2(\mathbb{A}^\infty)$ of weight (k, η). Then there is a CM field L_π and for each finite place λ of L_π a continuous irreducible Galois representation

$$\rho_\lambda(\pi) : G_F \to \mathrm{GL}_2(\overline{L}_{\pi,\lambda})$$

such that

(1) if π_v is unramified and v does note divide the residue characteristic of λ, then $\rho_\lambda(\pi)|_{G_{F_v}}$ is unramified, and the characteristic polynomial of Frob_v is

$$X^2 - t_v X + (\sharp k_v) s_v,$$

where T_v and S_v are the eigenvalues of T_v and S_v respectively on $\pi^{\mathrm{GL}_2(\mathcal{O}_{F_v})}$ (cf. Remark 3.7). By the Chebotarev density theorem, this already characterises $\rho_\lambda(\pi)$ up to isomorphism.

(2) the Frobenius semisimplification of the Weil-Deligne representation associated with $\rho_\lambda(\pi)|_{G_{F_v}}$ is isomorphic to $\mathrm{rec}_{F_v}(\pi_v \otimes |\det|^{-1/2})$.

(3) If v divides the residue characteristic of λ, then $\rho_\lambda(\pi)|_{G_{F_v}}$ is de Rham with τ-Hodge-Tate weights η_τ, $\eta_\tau + k_\tau - 1$, where $\tau : F \hookrightarrow \overline{L}_\pi \subset \mathbb{C}$ is an embedding lying over v. Moreover, if π_v is unramified, then $\rho_\lambda(\pi)|_{G_{F_v}}$ is crystalline.

(4) $\det r_\lambda(\pi)(c) = -1$ for each complex conjugation c.

Using the global Jacquet-Langlands correspondence, we may associate a Galois representation to the infinite dimensional cuspidal automorphic representations of $G_D(\mathbb{A}^\infty)$ for any D. On the other hand, in order to prove the existence of the Galois representation, we use the Jacquet-Langlands correspondence and transfer to a D over F for which $S(D)$ contains all but one infinite place. Then the Galois representations are realized in the étale cohomology group of the associated Shimura curve. The remaining Galois representations are constructed from these ones via congruences. [81, 16]

3.3.6. *Base change theory.* Fix E a cyclic extension of the number field F, of prime degree l. Roughly speaking, the theory of *base change* describes the correspondence between automorphic representations of the groups $\mathrm{GL}_n(\mathbb{A}_F)$ and $\mathrm{GL}_n(\mathbb{A}_E)$ which reflects the operations of *restriction* of Galois representations of W_F to W_E. The first results on base change for automorphic forms used the theory of L-functions, and were restricted to the case of quadratic E and GL_2. The introduction of the *trace formula* is due to H. Saito, who dealt with GL_2 and arbitrary cyclic E using the classical language of automorphic forms ([65]). Immediately after that, Shintani reformulated Saito's result using group representations, and gave the correct local definition of base change lifting ([74]). Finally, Langlands saw the connection with Artin's conjecture, and reshaped the trace formula proof for GL_2 in a form suitable for later generalizations to GL_n developed by Arthur and Clozel (cf. [55, 1]). Since only the case $n = 2$ is required here, we restricted ourselves to this case.

Let E/F be a cyclic extension of totally real fields of prime degree. Let $\mathrm{Gal}(E/F) = \langle\sigma\rangle$ and let $\mathrm{Gal}(E/F)^\vee = \langle\delta_{E/F}\rangle$ be the dual group of $\mathrm{Gal}(E/F)$. Let π be a cuspidal automorphic representation of $\mathrm{GL}_2(\mathbb{A}_F^\infty)$ of weight (k, η). Then there is a cuspidal automorphic representation $\mathrm{BC}_{E/F}(\pi)$ of $\mathrm{GL}_2(\mathbb{A}_E^\infty)$ of weight $(\mathrm{BC}_{E/F}(k), \mathrm{BC}_{E/F}(\eta))$ such that the following claims hold.

(1) Let v be a finite place of F and w be a finite place of E with $w|v$. Then $\mathrm{rec}_{E_w}(\mathrm{BC}(\pi)_w) = \mathrm{rec}_{F_v}(\pi_v)|_{W_{E_w}}$. In particular, $r_\lambda(\mathrm{BC}_{E/F}(\pi)) \cong r_\lambda(\pi)|_{G_E}$.
(2) $\mathrm{BC}_{E/F}(k)_w = k_v$ and $\mathrm{BC}_{E/F}(\eta)_w = \eta_v$.
(3) $\mathrm{BC}_{E/F}(\pi) \cong \mathrm{BC}_{E/F}(\pi')$ if and only if $\pi \cong \pi' \otimes (\delta_{E/F}^i \circ \mathrm{Art}_F \circ \det)$ for some i.
(4) A cuspidal automorphic representation Π of $\mathrm{GL}_2(\mathbb{A}_E^\infty)$ is in the image of $\mathrm{BC}_{E/F}$ if and only if $\Pi \circ \sigma \cong \Pi$.

Note that $\mathrm{BC}_{E/F}$ exists for general cyclic extension E/F and it satisfies similar properties as above except one point. In the general case, the representation $\mathrm{BC}_{E/F}(\pi)$ is automorphic. It is cuspidal (as opposed to just automorphic) unless E/F is quadratic over F, and π is monomial (i.e. $r_\lambda(\pi) = \mathrm{Ind}_{W_E}^{W_F} \theta$).

The theory of base change is a strong tool in the study of automprhic representations. In particular, it has two consequences that are important in the study of modularity representations: the Langlands-Tunnell theorem used in [92] and the simplification in Section 3.3.9.

3.3.7. *The Langlands-Tunnell theorem.* As a partial converse of Section 3.3.5, the Langlands-Tunnell theorem associated a cuspidal automorphic representation to a Galois representation. Suppose F is a number field and the irreducible representation

$$\rho : W_F \to \mathrm{GL}_2(\mathbb{C})$$

has a solvable image in $\mathrm{PGL}_2(\mathbb{C})$. Then there exists a (unique) irreducible cuspidal automorphic representation $\pi(\rho) = \otimes\pi_v$ of $\mathrm{GL}_2(\mathbb{A}_F^\infty)$ such that

$$\mathrm{Tr}(\rho(\mathrm{Frob}_v)) = \mathrm{Tr}(T_v)$$

for almost all v. (cf. [55, 87].)

3.3.8. *Modular Galois representations.*

Definition 3.11. A Galois representation $r : G_F \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$ is *modular* (of weight (k,η)) if it is isomorphic to $i(\rho_\lambda(\pi))$ for some cuspidal automorphic representation π (of weight (k,η)) and some $i : L_\pi \hookrightarrow \overline{\mathbb{Q}}_p$ lying over λ.

As an application of base change, we have the following result.

Proposition 3.12. *Suppose that $r : G_F \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$ is a continuous representation, and that E/F is a finite solvable Galois extension of totally real fields. Then $r|_{G_E}$ is modular if and only if r is modular.*

Proof. Without loss of generality, we may assume that E/F is cyclic of prime degree. If r is modular, then $r|_{G_E}$ is modular by base change. Conversely, suppose that $r|_{G_E}$ is modular, say $r|_{G_E} \cong i(r_\lambda(\Pi))$, where Π is a cuspidal automorphic representation of $\mathrm{GL}_2(\mathbb{A}_E)$. By multiplicity one result (cf. Section 3.3.1), we must have $\Pi \circ \sigma \cong \Pi$. Therefore, there is an automorphic representation π of $\mathrm{GL}_2(\mathbb{A}_F)$ such that $\mathrm{BC}_{E/F}(\pi) = \Pi$. As $i(r_\lambda(\pi))|_{G_E} = r|_{G_E}$ is irreducible, we must have $r \cong i(r_\lambda(\pi)) \otimes \chi \cong i(r_\lambda(\pi \otimes (\chi \circ \mathrm{Art}^{-1} \circ |\det|^{-1/2})))$ for some χ. Hence r is modular. □

3.3.9. *The simplification.* We need the following result from [83, Lemma 2.2].

Lemma 3.13. *Let K be a number field and let S be a finite set of places of K. For each $v \in S$, let L_v be a finite Galois extension of K_v. Then there is a finite solvable Galois extension M/K such that for each place w of M above a place $v \in S$ there is an isomorphism $L_v \cong M_w$ as K_v-algebras.*

In order to prove Theorem 1.1, by replacing F with a solvable totally real extension which is unramified at all primes above p, we may assume that

- $[F : \mathbb{Q}]$ is even.
- $\overline{\rho}$ is unramified outside p.
- For all primes $v \nmid p$, both $\rho(I_{F_v})$ and $\rho_0(I_{F_v})$ are unipotent (possible trivial).
- If ρ or ρ_0 are ramified at some place $v \nmid p$, then $\overline{\rho}|_{G_{F_v}}$ is trivial and $\sharp k(v) \equiv 1 \pmod{p}$.
- $\det \rho = \det \rho_0$. (To see this, note that the assumption that ρ and ρ_0 are crystalline with the same Hodge-Tate weights for all places dividing p implies that $\det \rho / \det \rho_0$ is unramified at all places dividing p. Since we have already assumed that $\rho(I_{F_v})$ and $\rho_0(I_{F_v})$ are unipotent for all $v \nmid p$, therefore $\det \rho / \det \rho_0$ is unramified at all

places, and thus has finite order. Moreover, it is residually trivial, it has p-power order and thus is trivial on all complex conjugations as $p \neq 1$. The extension cut out by the kernel of the character $\det \rho / \det \rho_0$ is a finite, abelian, totally real extension unramified at all places dividing p. The claim then follows by base change.)

From now on, we assume that all of these conditions hold. Write χ for $\det \rho = \det \rho_0$. Then $\chi \epsilon_p = \chi_{0,\iota}$ for some algebtaic grossencharacter χ_0. We assume further that the coefficient field L is sufficiently large, in the sense that L contains a primitive p-th root of unity, and for all $g \in G_F$, $\mathbb{F}$ contains the eigenvalues of $\overline{\rho}(g)$.

3.4. **The integral theory of automorphic forms over quaternion algebras.** In order to prove Theorem 1.1, we will need to study congruences between modular/ automorphic forms. In order to do so, it is convenient to work with automorphic forms on $G_D(\mathbb{A}^\infty)$ with $S(D) = S_\infty$. This is possible as $[F : \mathbb{Q}]$ is even. In this case, $G_D(\mathbb{A}^\infty) = \mathrm{GL}_2(\mathbb{A}^\infty)$ and $(D \otimes_{\mathbb{Q}} \mathbb{R})^\times / (F \otimes_{\mathbb{Q}} \mathbb{R})^\times$ is compact.

3.4.1. *Definition.* Fix an isomorphism $\iota : \overline{L} \xrightarrow{\sim} \mathbb{C}$ and some $k \in \mathbb{Z}_{\geq 2}^{\mathrm{Hom}(F,\mathbb{C})}$, $\eta \in \mathbb{Z}^{\mathrm{Hom}(F,\mathbb{C})}$ with $k_\tau + 2\eta_\tau - 1$ independent of $\tau \in \mathrm{Hom}(F, \mathbb{C})$. Denote this number by w. Let $U = \prod_v U_v \subset \mathrm{GL}_2(\mathbb{A}^\infty)$ be a open compact subgroup, and let S be a finite set of finite places of F, not containing any of the places over p, with the property that $U_v = \mathrm{GL}_2(\mathcal{O}_{F_v})$ if $v \notin S$.

Let $U_S := \prod_{v \in S} U_v$ and write $U = U_S U^S$. Let $\psi : U_S \to \mathcal{O}^\times$ be a continuous homomorphism and let $\chi_0 : \mathbb{A}_F^\times / F^\times \to \mathbb{C}^\times$ be an algebraic grossencharacter such that

- χ_0 is unramified outside S;
- for each place $v|\infty$, $\chi_0|_{(F_v^\times)^\circ}(x) = x^{1-w}$, where A° means the component of identity of A;
- $\chi_0|_{(\prod_{v \in S}) \cap U_S} = \psi^{-1}$.

Then χ_0 gives us a character

$$\chi_{0,\iota} : \mathbb{A}_F^\times / \overline{F^\times (F_\infty^\times)^\circ} \to \overline{L}^\times$$
$$x \mapsto \Big(\prod_{\tau : F \to L} \tau(x_p)^{1-w}\Big) \iota^{-1}\Big(\prod_{\tau : F \hookrightarrow \mathbb{C}} \tau(x_\infty)\Big)^{w-1} \chi_0(x).$$

The spaces of algebraic automorphic forms will be defined in a similar way to the classical spaces defined in Section 2.2, but with the role of the infinite places being played by the places lying over p (cf. [84, 50, 36]). First, we define coefficient systems as follows. Let

$$\Lambda = \Lambda_{k,\eta,\iota} = \bigotimes_{\tau : F \hookrightarrow \mathbb{C}} \mathrm{Sym}^{k_\tau - 2} \mathcal{O}^2 \otimes (\wedge^2 \mathcal{O}^2)^{\otimes \eta_\tau}.$$

Let $\mathrm{GL}_2(\mathcal{O}_{F,p}) := \prod_{v|p} \mathrm{GL}_2(\mathcal{O}_{F_v})$ act on Λ via $\iota^{-1}\tau$ on the τ-component. In particular,

$$\Lambda \otimes_{\iota:\mathcal{O}\to\mathbb{C}} \mathbb{C} \cong \bigotimes_{\tau:F\hookrightarrow\mathbb{C}} \mathrm{Sym}^{k_\tau-2}\,\mathbb{C}^2 \otimes (\wedge^2\mathbb{C}^2)^{\otimes\eta_\tau},$$

which has an obvious action of $\mathrm{GL}_2(F_\infty)$, and the two actions of $\mathrm{GL}_2(\mathcal{O}_{F,(p)})$ via its embeddings into $\mathrm{GL}_2(\mathcal{O}_{F,p})$ and $\mathrm{GL}_2(F_\infty)$ are compatible.

Let A be a finite $\mathcal{O}$-module. Define $S(U,A) = S_{k,\eta,\psi,\chi_0}(U,A)$ to be the space of functions

$$\phi : D^\times \backslash \mathrm{GL}_2(\mathbb{A}_F^\infty) \to \Lambda \otimes_{\mathcal{O}} A$$

such that

- $\phi(gu) = \psi(u_S)^{-1}u_p^{-1}\phi(g)$ for all $g \in \mathrm{GL}_2(\mathbb{A}_F^\infty)$ and $u \in U$;
- $\phi(gz) = \chi_{0,\iota}(z)\phi(g)$ for all $g \in \mathrm{GL}_2(\mathbb{A}_F^\infty)$ and $z \in (\mathbb{A}_F^\infty)^\times$.

Since $D^\times\backslash \mathrm{GL}_2(\mathbb{A}_F^\infty)/(\mathbb{A}_F^\infty)^\times U$ is finite, $S(U,A)$ is a finite free $\mathcal{O}$-module. More precisely, write

$$\mathrm{GL}_2(\mathbb{A}_F^\infty) = \coprod_{i\in I} D^\times g_i U (\mathbb{A}_F^\infty)^\times$$

for some finite indexing set I, we have an injection

$$\begin{aligned} S(U,A) &\hookrightarrow \oplus_{i\in I}(\Lambda\otimes A) \\ \phi &\mapsto (\phi(g_i)). \end{aligned}$$

To determine the image, we need to consider the equation $g_i = \delta g_i u z$ for $\delta \in D^\times$, $u \in U$, and $z \in (\mathbb{A}_F^\infty)^\times$, because then $\phi(g_i) = \phi(\delta g_i u z) = \chi_{0,\iota}(z)\psi(u_s)^{-1}u_p^{-1}\phi(g_i)$. From this we obtain an isomorphism

$$S(U,A) \xrightarrow{\sim} \oplus_{i\in I}(\Lambda\otimes A)^{(U(\mathbb{A}_F^\infty)^\times \cap g_i^{-1}D^\times g_i)/F^\times}. \tag{3.4}$$

Lemma 3.14. *Each group* $(U(\mathbb{A}_F^\infty)^\times \cap g_i^{-1}D^\times g_i)/F^\times$ *is finite. If* $p > 3$ *and* p *is unramified in* F*, the order of* $(U(\mathbb{A}_F^\infty)^\times \cap g_i^{-1}D^\times g_i)/F^\times$ *is not divisible by* p*.*

Proof. This is [84, Lemma 1.1]. Set $V = \prod_{v\nmid\infty}\mathcal{O}_{F,v}^\times$. Then we have exact sequences

$$0 \to (UV\cap g_i^{-1}D^{\det=1}g_i)/\{\pm1\} \to (U(\mathbb{A}_F^\infty)^\times \cap g_i^{-1}D^\times g_i)/F^\times \to (((\mathbb{A}_F^\infty)^\times)^2 V\cap F^\times)/(F^\times)^2$$

and

$$0 \to \mathcal{O}_F^\times/(\mathcal{O}_F^\times)^2 \to (((\mathbb{A}_F^\infty)^\times)^2 V \cap F^\times)/(F^\times)^2 \to H[2] \to 0,$$

where H is the class group of $\mathcal{O}_F$. We see that $(((\mathbb{A}_F^\infty)^\times)^2 V \cap F^\times)/(F^\times)^2$ is finite of 2-power order. Moreover $UV \cap g_i^{-1}D^{\det=1}g_i$ is finite. For $p > 3$ and p unramified in F, $D^\times$ and hence $UV \cap g_i^{-1}D^{\det=1}g_i$ contain no elements of order exactly p. The lemma follows. □

Remark 3.15. If U is sufficiently small so that $(UV \cap g_i^{-1}D^{\det=1}g_i)/\{\pm 1\}$ is trivial, then $(U(\mathbb{A}_F^\infty)^\times \cap g_i^{-1}D^\times g_i)/F^\times$ is a 2-group.

Remark 3.16. Let $g_i^{-1}\delta g_i$ represents an element in $(U(\mathbb{A}_F^\infty)^\times \cap g_i^{-1}D^\times g_i)/F^\times$ with $\delta \in D^\times$, we see that $\delta^2/\det\delta \in D^\times g_i U g_i^{-1}(\det U)$, the intersection of a discrete set and a compact set, so $\delta^2/\det\delta$ has finite order, i.e. is a root of unity. However any element of D generates an extension of F of degree at most 2, it must be a root of unity of degree prime to p since $[F(\zeta_p) : F] > 2$.

Corollary 3.17. *With the above notation, the following claims hold.*

(1) $S(U,\mathcal{O}) \otimes_{\mathcal{O}} A \xrightarrow{\sim} S(U,A)$.
(2) *If V is an open normal subgroup of U with $[U : V]$ a power of p, then $S(V,\mathcal{O})$ is a free $\mathcal{O}[U/V(U \cap (\mathbb{A}_F^\infty)^\times)]$-module.*

Proof. Denote by G_i the group $(U(\mathbb{A}_F^\infty)^\times \cap g_i^{-1}D^\times g_i)/F^\times$. Since G_i has order prime to p, we have $(\Lambda \otimes A)^{G_i} = \Lambda^{G_i} \otimes A$. The first claim follows immediately from the isomorphism $S(U,A) \xrightarrow{\sim} \oplus_{i\in I}(\Lambda \otimes A)^{G_i}$.

For the second claim, write $U = \coprod_{j\in J} u_j V(U \cap (\mathbb{A}_F^\infty)^\times)$. It suffices to prove that $\mathrm{GL}_2(\mathbb{A}_F^\infty) = \coprod_{i\in I, j\in J} D^\times g_i u_j V(\mathbb{A}_F^\infty)^\times$. To see this, we need to show that if $g_i u_j = \delta g_{i'} u_{j'} v z$ for some $\delta \in D^\times$, $v \in V$, $z \in (\mathbb{A}_F^\infty)^\times$, then $i = i'$ and $j = j'$.

From the definition of I, it is clear that $i = i'$. Then we have $u_{j'} v u_j^{-1} z = g_i^{-1}\delta^{-1} g_i$. The argument as in Remark 3.16 tells us that there is some positive integer N coprime to p such that $\delta^N \in F^\times$, so $(u_{j'} v u_j^{-1})^N \in (\mathbb{A}_F^\infty)^\times$. Since V is normal in U, we can write $(u_{j'} v u_j^{-1})^N = (u_{j'} u_j^{-1})^N v'$ for some $v' \in V$. Therefore $(u_{j'} u_j^{-1})^N \in V(U \cap (\mathbb{A}_F^\infty)^\times)$. Since $[U : V]$ is a power of p, we see that $u_{j'} u_j^{-1} \in V(U \cap (\mathbb{A}_F^\infty)^\times)$ and hence $j = j'$. ☐

3.4.2. *The Hecke algebra.* If $v \nmid p$ or if $v | p$ but $\tau|_{U_v} = 1$, then the Hecke algebra $\mathcal{O}[U_v \backslash \mathrm{GL}_2(F_v)/U_v]$ acts on $S(U,\mathcal{O})$. Explicitly, if

$$U_v h U_v = \coprod_i h_i U_v,$$

then

$$([U_v h U_v] f)(g) = \sum_i f(g h_i).$$

Let $\tilde{\mathbb{T}} := \mathcal{O}[T_v, S_v : \ v \nmid p, v \notin S]$, where T_v and S_v be the double coset operators corresponding to $\begin{pmatrix} \varpi_v & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} \varpi_v & 0 \\ 0 & \varpi_v \end{pmatrix}$ respectively. Let $\mathbb{T}_U$ be the image of $\tilde{\mathbb{T}}$ in $\mathrm{End}_{\mathcal{O}}(S(U,\mathcal{O}))$, so that $\mathbb{T}_U$ is a commutative $\mathcal{O}$-algebra which acts faithfully on $S(U,\mathcal{O})$ and finite free as an $\mathcal{O}$-module.

By definition, we have

$$S(U,\mathcal{O})\otimes_{\mathcal{O},\iota}\mathbb{C}\xrightarrow{\sim}\mathrm{Hom}_{U_S}(\mathbb{C}(\psi^{-1}),S_{k,\eta}^{U^S,\chi_0})$$
$$\phi\mapsto(g\mapsto g_\infty^{-1}\iota(g_p\phi(g^\infty))),$$

where the target of the isomorphism is the set of elements $\phi\in S_{k,\eta}$ with $z\phi=\chi_0(z)\phi$ for all $x\in(\mathbb{A}_F^\infty)^\times$, $u\phi=\psi(u_S)^{-1}\phi$ for all $u\in U$. This isomorphism is compatible with the action of $\tilde{\mathbb{T}}$ on each side. The target is isomorphic to

$$\bigoplus_\pi \mathrm{Hom}_{U_S}(\mathbb{C}(\psi^{-1}),\pi_S)\otimes(\otimes'_{v\notin S}\pi_v^{\mathrm{GL}_2(\mathcal{O}_{F_v})}),$$

where the sum is over the cuspidal automorphic representations π of $G_D(\mathbb{A}^\infty)$ of weight (k,η), which have central character χ_0 and are unramified outside S.

By strong multiplicity one (cf. Section 3.3.1), we have an isomorphism

$$\mathbb{T}_U\otimes_{\mathcal{O},\iota}\mathbb{C}\cong\prod_{\pi\text{ as above, with }\mathrm{Hom}_{U_S}(\mathbb{C}(\psi^{-1}),\pi_S)\neq 0}\mathbb{C}$$

sending T_v, S_v to their eigenvalues on $\pi_v^{\mathrm{GL}_2(\mathcal{O}_{F_v})}$. In particular $\mathbb{T}_U$ is reduced. Moreover, this shows that there is a bijection between ι-linear ring homomorphisms $\theta:\mathbb{T}_U\to\mathbb{C}$ and the set of π as above, where π corresponds to the character taking T_v, S_v to their corresponding eigenvalues.

As explained in Section 3.3.5, each π has a corresponding Galois representation. Taking the product of these represntations, we obtain a representation

$$\rho^{\mathrm{mod}}:G_F\to\prod_\pi\mathrm{GL}_2(\overline{L})=\mathrm{GL}_2(\mathbb{T}_U\otimes_{\mathcal{O}}\overline{L}),$$

which is characterized by the properties that it is unramified outside $S\cup\{v:\ v|p\}$, and for any unramified v, $\rho^{\mathrm{mod}}(\mathrm{Frob}_v)$ has characteristic polynomial

$$X^2-T_vX+(\sharp k_v)S_v=0.$$

Let $\mathfrak{m}$ be a maximal ideal of $\mathbb{T}_U$. Then if $\mathfrak{p}\subsetneq\mathfrak{m}$ is a minimal prime, then there is an injection $\theta:\mathbb{T}_U/\mathfrak{p}\hookrightarrow\overline{L}$, which corresponds to some π above. The semisimple mod p Galois representation corresponding to π can be conjugated to give a representation $\overline{\rho}_{\mathfrak{m}}:G_F\to\mathrm{GL}_2(\mathbb{T}_U/\mathfrak{m})$. This is well defined up to isomorphism independently of the choice of $\mathfrak{p}$ and θ.

Since $\mathbb{T}_U$ is finite over the complete local ring $\mathcal{O}$, it is semilocal. Write $\mathbb{T}_U=\prod_{\mathfrak{m}}\mathbb{T}_{U,\mathfrak{m}}$. Suppose that $\overline{\rho}_{\mathfrak{m}}$ is absolutely irreducible. Then we have the representation

$$\rho^{\mathrm{mod}}:G_F\to\mathrm{GL}_2(\mathbb{T}_{U,\mathfrak{m}}\otimes_{\mathcal{O}}\overline{L})=\prod_\pi\mathrm{GL}_2(\overline{L}),$$

where the product is over the π as above with $\overline{\rho}_{\pi,\iota} \cong \overline{\rho}_{\mathfrak{m}}$. Each representation to $\mathrm{GL}_2(\overline{L})$ can be conjugated to lie in $\mathrm{GL}_2(\mathcal{O}_{\overline{L}})$, and after further conjugation (so that the residual representations are equal to $\overline{\rho}_{\mathfrak{m}}$, rather than just conjugate to it), the image of $\rho_{\mathfrak{m}}^{\mathrm{mod}}$ lies in the subring of $\prod_\pi \mathrm{GL}_2(\mathcal{O}_{\overline{L}})$ consisting of elements whose image modulo the maximal ideal of $\mathcal{O}_{\overline{L}}$ lie in $\mathbb{T}_U/\mathfrak{m}$. Then $\rho_{\mathfrak{m}}^{\mathrm{mod}}$ can be conjugated to lie in $\mathrm{GL}_2(\mathbb{T}_{U,\mathfrak{m}})$. We will write $\rho_{\mathfrak{m}}^{\mathrm{mod}} : G_F \to \mathrm{GL}_2(\mathbb{T}_{U,\mathfrak{m}})$ for the resulting representation from now on.

Remark 3.18. Later in the notes we need to consider Hecke operators at places in S. To this end, let $T \subset S$ such that $\psi|_{U_T} = 1$ and choose $g_v \in \mathrm{GL}_2(F_v)$ for each $v \in T$. Set $W_v = [U_v g_v U_v]$ and define $\tilde{\mathbb{T}}'_U = \tilde{\mathbb{T}}_U[W_v : v \in T]$. Similarly we have $\mathbb{T}_U \subset \mathbb{T}'_U \subset \mathrm{End}_{\mathcal{O}}(S(U,\mathcal{O}))$. The algebra $\mathbb{T}'_U$ is commutative, and finite and flat over $\mathcal{O}$. However it need not be reduced. Indeed, we have

$$\mathbb{T}'_U \otimes_{\mathcal{O},\iota} \mathbb{C} \cong \bigoplus_\pi \otimes_{v\in T}\{\text{subalgebra of } \mathrm{End}_{\mathbb{C}}(\pi_v^{U_v}) \text{ generated by } W_v\},$$

so that there is a bijection between ι-linear homomorphisms $\mathbb{T}'_U \to \mathbb{C}$ and tuples $(\pi, \{\alpha_v\}_{v\in T})$, where α_v is an eigenvalue of W_v on $\pi_v^{U_v}$.

4. The Taylor-Wiles-Kisin method

In this section, we review the patching result of Kisin [50, Proposition 3.3.1]. It is a criterion for a map of $\mathcal{O}$-algebras to be isomorphism up to p-torsion. This will be applied to establish that certain Galois deformation rings and Hecke rings are isomorphic up to p-torsion. The argument is analogous to that of Taylor-Wiles [86] and Diamond [24]. One of the differences with this approach is that Kisin's criterion could be used to treat both the minimal and non-minimal case.

Proposition 4.1. *Let B be a complete local, flat $\mathcal{O}$-algebra, which is a domain of dimension $d+1$, and such that $B[1/p]$ is formally smooth over E. Suppose that R is a B-algebra and M is a non-zero R-module, and that there are non-negative integers h and j such that for each non-negative integer n there are maps of $\mathcal{O}$-algebras*

$$\mathcal{O}[[y_1, \dots, y_{h+j}]] \to R_n \to R \tag{4.1}$$

and a map of R_n-modules $M_n \to M$ satisfying the following conditions:

1. *The maps $R_n \to R$ and $M_n \to M$ are surjective, and the first is a map of B-algebras.*
2. *$(y_1, \dots, y_h)R_n = \mathrm{Ker}(R_n \to R)$ and $(y_1, \dots, y_h)M_n = \mathrm{Ker}(M_n \to M)$.*

(3) *If* $\mathfrak{b}_n \subset \mathcal{O}[[y_1, \ldots, y_{h+j}]]$ *is the annihilator of* M_n*, then*

$$\mathfrak{b}_n \subset ((1+y_1)^{p^n} - 1, \ldots, (1+y_h)^{p^n} - 1),$$

and M_n *is finite free over* $\mathcal{O}[[y_1, \ldots, y_{h+j}]]/\mathfrak{b}_n$*. So, in particular,* M *is finite free over* $\mathcal{O}[[y_{h+1}, \ldots, y_{h+j}]]$*.*

(4) R_n *is a quotient of* $B[[x_1, \ldots, x_{h+j-d}]]$*.*

Then the R *is a finite* $\mathcal{O}[[y_{h+1}, \ldots, y_{h+j}]]$*-algebra, and* $M \otimes_{\mathcal{O}} E$ *is a finite projective and faithful* $R[1/p]$*-module.*

To apply this result, R is certain universal deformation ring and M is the space of certain automorphic forms. As the action of R on M comes form the surjection $R \twoheadrightarrow T$, where T is certain Hecke algebra, the faithfulness statement implies that $R[1/p] = T[1/p]$.

Proof. For a complete local ring A, denote by $\mathfrak{m}_A$ its maximal ideal. For a non-negative integer n, denote by $\mathfrak{m}_A^{(n)} \subset \mathfrak{m}_A$ the ideal generated by the elements of $\mathfrak{m}_A$ which are n-th powers.

Let s denote the $\mathcal{O}[[y_{h+1}, \ldots, y_{h+j}]]$-rank of M. This is also the $\mathcal{O}[[y_1, \ldots, y_{h+j}]]/\mathfrak{b}_n$-rank of M_n. For a non-negative integer m, write $r_m = smp^m(h+j)$, and

$$\mathfrak{c}_m = (\pi_E^m, (y_1+1)^{p^m} - 1, \ldots, (y_h+1)^{p^m} - 1, y_{h+1}^{p^m}, \ldots, y_{h+j}^{p^m}) \subset \mathcal{O}[[y_1, \ldots, y_{h+1j}]].$$

For $m \geq 1$ a *patching datum* (D, L) of level m consists of

(1) A sequence of maps of $\mathcal{O}$-algebras

$$\mathcal{O}[[y_1, \ldots, y_{h+j}]]/\mathfrak{c}_m \to D \to R/(\mathfrak{c}_m R + \mathfrak{m}_R^{(r_m)}) \tag{4.2}$$

where the second map is a surjective map of B-algebras and $\mathfrak{m}_D^{(r_m)} = 0$.

(2) A surjection of B-algebras $B[[x_1, \ldots, x_{h+j-d}]] \to D$.

(3) A D-module L which is finite free over $\mathcal{O}[[y_1, \ldots, y_{h+j}]]/\mathfrak{c}_m$ of rank s, and a surjection of D-modules $L \to M/\mathfrak{c}_m M$.

A morphism of patching data $(D_1, L_1) \to (D_2, L_2)$ consists of a pair of morphisms $\alpha : D_1 \to D_2$ and $\beta : L_1 \to L_2$ such that:

(1) $\alpha : D_1 \to D_2$ is a map of $B[[x_1, \ldots, x_{h+j-d}]]$-algebras and is compatible with (4.2), i.e., the following diagram is commutative

$$\begin{array}{ccccc}
\mathcal{O}[[y_1, \ldots, y_{h+j}]]/\mathfrak{c}_m & \longrightarrow & D_1 & \longrightarrow & R/(\mathfrak{c}_m R + \mathfrak{m}_R^{(r_m)}) \\
\| & & \downarrow{\scriptstyle \alpha} & & \| \\
\mathcal{O}[[y_1, \ldots, y_{h+j}]]/\mathfrak{c}_m & \longrightarrow & D_2 & \longrightarrow & R/(\mathfrak{c}_m R + \mathfrak{m}_R^{(r_m)})
\end{array}$$

(2) $\beta : L_1 \to L_2$ is a surjection of D_1-modules which is compatible with the surjections of L_1 and L_2 onto $M/\mathfrak{c}_m M$.

Since the number of elements of D is bounded by $B[[x_1, \dots, x_{h+j-d}]]/\mathfrak{m}^{(r_m)}_{B[[x_1,\dots,x_{h+j-d}]]}$, there are only finitely many isomorphism classes of patching data.

Given positive integers n and m with $n \geq m$, we define a patching datum $(D_{m,n}, L_{m,n})$ of level m, by taking

$$D_{m,n} = R_n/(\mathfrak{c}_m R_n + \mathfrak{m}^{(r_m)}_{R_n}) \text{ and } L_{m,n} = M_n/\mathfrak{c}_m M_n.$$

To check that $L_{m,n}$ is a $D_{m,n}$-module we have to show that $\mathfrak{m}^{(r_m)}_{R_n} M_n \subset \mathfrak{c}_m M_n$. To see this, let $a \in \mathfrak{m}_{R_n}$. Then a induces a nilpotent endomorphism of $M/(\pi_E, y_{h+1}, \dots, y_{h+j})M$, so a^s induces the zero endomorphism. Hence $a^s M_n \subset (\pi_E, y_1, \dots, y_{h+j})M_n$, and

$$\begin{aligned} a^{sp^m(h+j)} M_n \subset &(\pi_E, y_1^{p^m}, \dots, y_{h+j}^{p^m})M \\ &= (\pi_E, (y_1+1)^{p^m} - 1, \dots, (y_h+1)^{p^m} - 1, y_{h+1}^{p^m}, \dots, y_{h+j}^{p^m})M_n. \end{aligned}$$

Finally $a^{r_m} M_n = a^{sp^m(h+j)m} M_n \subset \mathfrak{c}_m M_n$.

Since there are only finitely many isomorphism classes of patching data of level m, after replacing the sequence (R_n, M_n) by a subsequence, we may assume that for each $m \geq 1$, and $n \geq m$, the datum $(D_{m,n}, L_{m,n})$ is equal to $(D_{m,m}, L_{m,m})$. In particular, we have maps of patching data

$$(D_{m+1,m+1}, L_{m+1,m+1}) \to (D_{m,m}, L_{m,m}),$$

an isomorphism of $B[[x_1, \dots, x_{h+j-d}]]$-algebras

$$D_{m+1,m+1}/(\mathfrak{c}_m D_{m+1,m+1} + \mathfrak{m}^{(r_m)}_{D_{m+1,m+1}}) \xrightarrow{\sim} D_{m,m},$$

and an isomorphism of $D_{m+1,m+1}$-modules

$$L_{m+1,m+1}/\mathfrak{c}_m L_{m+1,m+1} \xrightarrow{\sim} L_{m,m}.$$

Now set $R_\infty = \varprojlim D_{m,m}$ and $M_\infty = \varprojlim L_{m,m}$. By construction we have a surjection

$$B[[x_1, \dots, x_{h+j-d}]] \to R_\infty,$$

and maps of complete local $\mathcal{O}$-algebras

$$\mathcal{O}[[y_1, \dots, y_{h+j}]] \to R_\infty \to R$$

where the second map is a map of B-algebras, and identifies $R_\infty/(y_1, \dots, y_h)R_\infty$ with R. We also have that M_∞ is a finite free $\mathcal{O}[[y_1, \dots, y_{h+j}]]$-module of rank $s > 0$, and that the R_n-module $M_\infty/(y_1, \dots, y_h)M_\infty$ is isomorphic to M. Note that

$$\dim \mathcal{O}[[y_1, \dots, y_{h+j}]] = h + j + 1 = d + 1 + h + j - d = \dim B[[x_1, \dots, x_{h+j-d}]].$$

By Lemma 4.2 below, $B[[x_1,\dots,x_{h+j-d}]]$ is a finite $\mathcal{O}[[y_1,\dots,y_{h+j}]]$-module, and $M_\infty \otimes_{\mathcal{O}} E$ is a finite projective, faithful $B[[x_1,\dots,x_{h+j-d}]][1/p]$-module. In particular we see that

$$B[[x_1,\dots,x_{h+j-d}]] \xrightarrow{\sim} R_\infty,$$

so that

$$M \otimes_{\mathcal{O}} E = M_\infty \otimes_{\mathcal{O}} E/(y_1,\dots,y_h)(M_\infty \otimes_{\mathcal{O}} E)$$

is a finite projective, faithful $R[1/p] = R_\infty[1/p]/(y_1,\dots,y_h)R_\infty[1/p]$-module, and R is finite over $\mathcal{O}[[y_{h+1},\dots,y_{h+j}]]$. The proposition follows. □

Lemma 4.2. *Let $A \xrightarrow{\varphi} D$ be a map of Noetherian domains of the same finite dimension d, and L a non-zero D-module which is finite and projective over A. Then φ is a finite map. If A and D are regular then L is a finite projective, faithful D-module.*

Proof. Let D' be the image of D in $\operatorname{End}_A L$. Then D' is finite over A, since L is finite over A. As L is a faithful A-module, so is D'. It follows that $\dim D' \geq d$, so that $D = D'$.

To show the second statement, we first remark that since A is a domain, L has the same rank $s > 0$ at all points of A. Similarly, if L is finite projective over D, it is a faithful D-module. Let $\mathfrak{p}$ be a prime of A, $\mathfrak{q}$ a prime of D lying over $\mathfrak{p}$. Let $\hat{A}_{\mathfrak{p}}$ and $\hat{D}_{\mathfrak{q}}$ be the completions of A and D at $\mathfrak{p}$ and $\mathfrak{q}$ respectively. It suffices to show that $L \otimes_D \hat{D}_{\mathfrak{q}}$ is a finite free $\hat{D}_{\mathfrak{q}}$-module. Thus we may replace A, D, and L by $\hat{A}_{\mathfrak{p}}$, $\hat{D}_{\mathfrak{q}}$, and $L \otimes_D \hat{D}_{\mathfrak{q}}$ respectively (note that $\hat{D}_{\mathfrak{q}}$ is finite over $\hat{A}_{\mathfrak{p}}$, and $L \otimes_D \hat{D}_{\mathfrak{q}}$ is a $\hat{A}_{\mathfrak{p}}$-direct summand of $L \otimes_A \hat{A}_{\mathfrak{p}}$), and assume that A and D are complete local regular rings.

Now the A-depth of L is d since L is A-free, hence the D-depth of L is $\geq d$, and therefore equal to d. The Auslander-Buchsbaum theorem then implies that L is D-free. (cf. [24, Theorem 2.1]). □

5. Patching and the proof of Theorem 1.1

As explained in Section 3.3.9, to prove Theorem 1.1, we may assume that the assumptions in Section 3.3.9 hold. Let D be a quaternion algebra over F ramified at exactly the infinite places. We work with automorphic representations of $G_D(\mathbb{A}_F^\infty)$.

Let T_p be the set of places of F lying over p, let T_r be the set of primes not lying over p at which ρ or ρ_0 is ramified, let $T = T_p \coprod T_r$. If $v \in T_r$, write σ_v for a choice of topological generator of I_{F_v}/P_{F_v}. By our assumptions, if $v \in T_r$, then $\overline{\rho}|_{G_{F_v}}$ is trivial, $\rho|_{I_{F_v}}$ and $\rho_0|_{I_{F_v}}$ are unipotent, and $\sharp(k(v)) \equiv 1 \pmod{p}$.

Let Q be a set of finite places of F such that, if $v \in Q$, then

- $v \notin T$;

- $\sharp k(v) \equiv 1 \pmod p$;
- $\overline{\rho}(\mathrm{Frob}_v)$ has distinct eigenvalues, which we denote $\overline{\alpha}_v$ and $\overline{\beta}_v$.

The existence of Q will the proved in Section 5.3.

5.1. **The deformation rings.** For each set Q of primes satisfying the above conditions, we define two deformation problems $\mathcal{S}_Q = (T \cup Q, \{\mathcal{D}_v\}, \chi)$ and $\mathcal{S}_Q = (T \cup Q, \{\mathcal{D}'_v\}, \chi)$ as follows. Let ζ be a fixed primitive p-th root of unity in L.

- If $v \in T_p$, then $\mathcal{D}_v = \mathcal{D}'_v$ is chosen so that
$$R^{\square}_{\overline{\rho}|_{G_{F_v}},\chi}/I(\mathcal{D}_v) = R^{\square}_{\overline{\rho}|_{G_{F_v}},\chi,\mathrm{cr},\{\mathrm{HT}_\sigma(\rho)\}},$$
i.e., we consider crystalline deformations of $\overline{\rho}|_{G_{F_v}}$ with determinant χ and with the same Hodge-Tate weights as ρ.
- If $v \in Q$, then $\mathcal{D}_v = \mathcal{D}'_v$ consists of all lifts of $\overline{\rho}|_{G_{F_v}}$ with determinant χ. In particular, we allow our deformations to ramify at places in Q.
- If $v \in T_r$, then $\mathcal{D}_v$ consists of all lifts of $\overline{\rho}|_{G_{F_v}}$ with $\mathrm{Char}_{\rho(\sigma_v)}(X) = (X-1)^2$, while $\mathcal{D}'_v$ consists of all lifts of $\overline{\rho}|_{G_{F_v}}$ with $\mathrm{Char}_{\rho(\sigma_v)}(X) = (X-\zeta)(X-\zeta^{-1})$. Note that the difference disappears if we modular λ since $\zeta \equiv 1 \pmod \lambda$.

We write
$$R^{\mathrm{loc}} = \hat{\otimes}_{v\in T,\mathcal{O}} R^{\square}_{\overline{\rho}|_{G_{F_v}},\chi}/I(\mathcal{D}_v), \quad R^{\mathrm{loc},'} = \hat{\otimes}_{v\in T,\mathcal{O}} R^{\square}_{\overline{\rho}|_{G_{F_v}},\chi}/I(\mathcal{D}'_v).$$

From our construction and the results proved on Galois deformations (cf. [93]), the following properties hold.

- $R^{\mathrm{loc}}/\lambda = R^{\mathrm{loc},'}/\lambda$.
- $(R^{\mathrm{loc},'})^{\mathrm{red}}$ is irreducible, $\mathcal{O}$-flat, and has Krull dimension $1 + 3\sharp T + [F:\mathbb{Q}]$.
- $(R^{\mathrm{loc}})^{\mathrm{red}}$ is $\mathcal{O}$-flat, equidimensional of Krull dimension $1 + 3\sharp T + [F:\mathbb{Q}]$, and the reduction modulo λ gives a bijection between the irreducible components of $\mathrm{Spec}\, R^{\mathrm{loc}}$ and those of $\mathrm{Spec}\, R^{\mathrm{loc}}/\lambda$.

We have the global analogues
$$R^{\mathrm{univ}}_Q := R^{\mathrm{univ}}_{\overline{\rho},\mathcal{S}_Q};\ R^{\mathrm{univ}'}_Q := R^{\mathrm{univ}}_{\overline{\rho},\mathcal{S}'_Q};\ R^{\square}_Q := R^{\square_T}_{\overline{\rho},\mathcal{S}_Q};\ R^{\square,'}_Q := R^{\square_T}_{\overline{\rho},\mathcal{S}'_Q}.$$

It is easy to see that the following properties hold.

- $R^{\mathrm{univ}}_Q/\lambda = R^{\mathrm{univ},'}_Q/\lambda$, and $R^{\square}_Q/\lambda = R^{\square,'}_Q/\lambda$.
- There are obvious natural maps $R^{\mathrm{loc}} \to R^{\square}_Q$, $R^{\mathrm{loc},'} \to R^{\square,'}_Q$, and these maps agree after reduction modulo λ.

We may and do fix representatives ρ_Q^{univ} and $\rho_Q^{\mathrm{univ},\prime}$ for the universal deformations of $\overline{\rho}$ over R_Q^{univ} and $R_Q^{\mathrm{univ},\prime}$ respectively, which are compatible with the choices of $\rho_\emptyset^{\mathrm{univ}}$ and $\rho_\emptyset^{\mathrm{univ},\prime}$, so that the induced surjections

$$R_Q^{\mathrm{univ}} \twoheadrightarrow R_\emptyset^{\mathrm{univ}}, \quad R_Q^{\mathrm{univ},\prime} \twoheadrightarrow R_\emptyset^{\mathrm{univ},\prime}$$

are identified modulo λ.

Fix a place $v_0 \in T$ and set $\mathcal{J} := \mathcal{O}[[X_{v,i,j}]]_{v\in T,\ i,j=1,2}/(X_{v_0,1,1})$. Let $\mathfrak{a}$ be the ideal of $\mathcal{J}$ generated by the $X_{v,i,j}$. Then our choice of ρ_Q^{univ} gives an identification $R_Q^{\square} \xrightarrow{\sim} R_Q^{\mathrm{univ}} \hat{\otimes}_{\mathcal{O}} \mathcal{J}$, corresponding to the universal T-framed deformation $(\rho_Q^{\mathrm{univ}}, \{1 + (X_{v,i,j})\}_{v\in T})$.

From the computation we made in [93], for each place $v \in Q$, we have an isomorphism

$$\rho_Q^{\mathrm{univ}}|_{G_{F_v}} \cong \chi_\alpha \oplus \chi_\beta,$$

where $\chi_\alpha,\ \chi_\beta : G_{F_v} \to (R_Q^{\mathrm{univ}})^\times$ are characters with $\chi_\alpha(\mathrm{Frob}_v) \equiv \alpha_v \pmod{\mathfrak{m}_{R_Q^{\mathrm{univ}}}}$ and $\chi_\beta(\mathrm{Frob}_v) \equiv \beta_v \pmod{\mathfrak{m}_{R_Q^{\mathrm{univ}}}}$.

Let Δ_v be the maximal p-power quotient of $k(v)^\times$ (which we sometimes regard as a subgroup of $k(v)^\times$). Then $\chi_\alpha|_{I_{F_v}}$ factors through the composite

$$I_{F_v} \twoheadrightarrow I_{F_v}/P_{F_v} \twoheadrightarrow k(v)^\times \twoheadrightarrow \Delta_v, \tag{5.1}$$

and if we write $\Delta_Q = \prod_{v\in Q} \Delta_v$, $(\prod_{v\in Q} \chi_\alpha) : \Delta_Q \to (R_Q^{\mathrm{univ}})^\times$, then we see that

$$(R_Q^{\mathrm{univ}})_{\Delta_Q} = R_\emptyset^{\mathrm{univ}}.$$

5.2. **The spaces of modular forms.** Recall we have fixed an isomorphism $\iota : \overline{L} \to \mathbb{C}$, and an algebraic grossencharacter χ_0 such that $\chi\epsilon_p = \chi_{0,\iota}$. Define k, η by $\mathrm{HT}_\tau(\rho_0) = \{\eta_{\iota\tau}, \eta_{\iota\tau} + k_{\iota\tau} - 1\}$. We define open compact subgroups $U_Q = \prod U_{Q,v}$, where

- $U_{Q,v} = \mathrm{GL}_2(\mathcal{O}_{F_v})$ if $v \notin Q \cup T_r$;
- $U_{Q,v} = U_0(v) = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathcal{O}_{F_v}) \mid c \equiv 0 \pmod{v}\}$ if $v \in T_r$;
- $U_{Q,v} = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in U_0(v) \mid ad^{-1} \pmod{v} \in k(v)^\times \mapsto 1 \in \Delta_v\}$ if $v \in Q$.

Let $\psi : \prod_{v\in T_r} U_{Q,v} \to \mathcal{O}^\times$ be the trivial character. Similarly, we set $U_Q' = U_Q$ and define $\psi' : \prod_{v\in T_r} U_{Q,v} \to \mathcal{O}^\times$ as follows. For each $v \in T_r$, we have a homomorphism $U_{Q,v} \to k(v)^\times$ given by sending $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ to $ad^{-1} \pmod{v}$, and we compose these characters with the characters $k(v)^\times \to \mathcal{O}^\times$ sending the image of σ_v to ζ, where σ_v is a generator of I_{F_v}/P_{F_v} (cf. equation (5.1)).

We obtain spaces of modular forms $S(U_Q, \mathcal{O})$, $S(U'_Q, \mathcal{O})$ and the corresponding Hecke algebras $\mathbb{T}_{U_Q}$ and $\mathbb{T}_{U_{Q'}}$. Here the Hecke algebras are generated by the Hecke operators T_v, S_v with $v \notin T \cup Q$, together with Hecke operators U_{ϖ_v} for $v \in Q$ defined by

$$U_{\varpi_v} = [U_{Q,v} \begin{pmatrix} \varpi_v & 0 \\ 0 & 1 \end{pmatrix} U_{Q,v}].$$

Note that $\psi \equiv \psi' \pmod{\lambda}$, so $S(U_\emptyset, \mathcal{O})/\lambda = S(U'_\emptyset, \mathcal{O})/\lambda$. Let $\mathfrak{m}_\emptyset \subset \mathbb{T}_{U_\emptyset}$ be the ideal generated by

- λ;
- $\operatorname{Tr} \overline{\rho}(\mathrm{Frob}_v) - Tv$, $v \notin T$;
- $\det \overline{\rho}(\mathrm{Frob}_v) - \sharp k(v) S_v$, $v \notin T$.

This is a proper maximal ideal of $\mathbb{T}_{U_\emptyset}$ as it is the kernel of the homomorphism $\mathbb{T}_{U_\emptyset} \to \mathcal{O} \twoheadrightarrow \mathbb{F}$, where $\mathbb{T}_{U_\emptyset} \to \mathcal{O}$ is the map coming from the automorphicity of ρ_0, sending T_v to $\operatorname{Tr} \rho_0(\mathrm{Frob}_v)$ and S_v to $\sharp k(v)^{-1} \det \rho_0(\mathrm{Frob}_v)$.

The universal property of $R_\emptyset^{\mathrm{univ}}$ gives us a surjection $R_\emptyset^{\mathrm{univ}} \twoheadrightarrow \mathbb{T}_\emptyset := \mathbb{T}_{U_\emptyset, \mathfrak{m}_\emptyset}$ and a corresponding lifting $\rho^{\mathrm{mod}} : G_F \to \mathrm{GL}_2(\mathbb{T}_\emptyset)$ of type $\mathcal{S}_\emptyset$. Similarly, we have a surjection $R_\emptyset^{\mathrm{univ},'} \twoheadrightarrow \mathbb{T}'_\emptyset := \mathbb{T}_{U'_\emptyset, \mathfrak{m}_\emptyset}$. Set $S_\emptyset := S(U_\emptyset, \mathcal{O})_{\mathfrak{m}_\emptyset}$ and $S'_\emptyset := S(U'_\emptyset, \mathcal{O})_{\mathfrak{m}_\emptyset}$. Then the identification $R_\emptyset^{\mathrm{univ}}/\lambda \cong R_\emptyset^{\mathrm{univ},'}/\lambda$ is compatible with $S_\emptyset/\lambda = S'_\emptyset/\lambda$. An important observation is the following result.

Lemma 5.1. *If* $\mathrm{Supp}_{R_\emptyset^{\mathrm{univ}}}(S_\emptyset) = \operatorname{Spec} R_\emptyset^{\mathrm{univ}}$*, then* ρ *is modular.*

Proof. Since $S_\emptyset$ is a faithful $\mathbb{T}_\emptyset$-module by definition, we see that $\mathrm{Ker}(R_\emptyset^{\mathrm{univ}} \to \mathbb{T}_\emptyset)$ is nilpotent, so that $(R_\emptyset^{\mathrm{univ}})^{\mathrm{red}} \xrightarrow{\sim} \mathbb{T}_\emptyset$. Then ρ corresponds to some homomorphism $R_\emptyset^{\mathrm{univ}} \to \mathcal{O}$, and thus to a homomorphism $\mathbb{T}_\emptyset \to \mathcal{O}$, and the composite of this homomorphism with $\iota : \mathcal{O} \to \mathbb{C}$ correponds to a cuspidal automorphic representation π of $G_D(\mathbb{A}_F^\infty)$ of weight (k, η), which by construction has the property that $\rho \cong \rho_{\pi,\iota}$, as required. $\square$

In order to apply Kisin's patching argument, we study the above constructions as Q varies. Let $\mathfrak{m}_Q$ be the maximal ideal of $\mathbb{T}_{U_Q}$ generated by

- λ;
- $\operatorname{Tr} \overline{\rho}(\mathrm{Frob}_v) - Tv$, $v \notin T \cup Q$;
- $\det \overline{\rho}(\mathrm{Frob}_v) - \sharp k(v) S_v$, $v \notin T \cup Q$;
- $U_{\varpi_v} - \overline{\alpha}_v$, $v \in Q$.

Write $S_Q = S_{U_Q} := S(U_Q, \mathcal{O})_{\mathfrak{m}_Q}$ and $\mathbb{T}_Q := (\mathbb{T}_{U_Q})_{\mathfrak{m}_Q}$. We have two homomorphisms $\Delta_Q \to \mathrm{End}(S_Q)$, one is given by

$$\delta \in \Delta_v \mapsto \begin{pmatrix} \delta & 0 \\ 0 & 1 \end{pmatrix},$$

and the other one is given by the composite

$$\Delta_Q \to R_Q^{\mathrm{univ}} \twoheadrightarrow \mathbb{T}_Q \to \mathrm{End}(S_Q).$$

5.2.1. *The local-global compatibility at places in Q.* A homomorphism $\theta : \mathbb{T}_Q \to \mathbb{C}$ corresponds to a cuspidal automorphic representation π, and for each $v \in Q$ the image α_v of U_{ϖ_v} is such that α_v is an eigenvalue of U_{ϖ_v} on $\pi_v^{U_{Q,v}}$.

It can be checked that since $\pi_v^{U_{Q,v}} \neq 0$, π_v is necessarily a subquotient of $I(\chi_1, \chi_2)$ for some tamely ramified characters $\chi_1, \chi_2 : F_v^\times \to \mathbb{C}^\times$. Then one checks explicitly that

$$(I(\chi_1, \chi_2))^{U_{Q,v}} \cong \mathbb{C}\phi_1 \oplus \mathbb{C}\phi_w,$$

where $w = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\phi_1(1) = \phi_w(w) = 1$, and $\mathrm{Supp}\,\phi_1 = B(F_v)U_{Q,v}$, $\mathrm{Supp}\,\phi_w = B(F_v)wU_{Q,v}$. As computed in [84, Section 2],

$$\begin{cases} U_{\varpi_v}\phi_1 &= \sharp k(v)^{1/2}\chi_1(\varpi_v)\phi_1 + X\phi_w \\ U_{\varpi_v}\phi_w &= \sharp k(v)^{1/2}\chi_2(\varpi_v)\phi_w. \end{cases} \tag{5.2}$$

Here $X = 0$ if $\chi_1\chi_2^{-1}$ is ramified. By local-global compatibility, $\iota^{-1}(\sharp k(v)^{1/2}\chi_1(\varpi_v))$ and $\iota(\sharp k(v)^{1/2}\chi_2(\varpi_v))$ are the eigenvalues of $\rho_{\pi,\iota}(\mathrm{Frob}_v)$, so one of them is a lift of $\overline{\alpha}_v$, and one is a lift of $\overline{\beta}_v$.

Moreover, easy computation shows that

$$\begin{pmatrix} \delta & 0 \\ 0 & 1 \end{pmatrix}\phi_1 = \chi_1(\delta)\phi_1, \quad \begin{pmatrix} \delta & 0 \\ 0 & 1 \end{pmatrix}\phi_w = \chi_2(\delta)\phi_w.$$

By local-global compatibility,

$$\rho_{\pi,\iota}|_{W_{F_v}}^{\mathrm{ss}} \cong (\chi_1|\cdot|^{-1/2} \oplus \chi_2|\cdot|^{-1/2}) \circ \mathrm{Art}_{F_v}^{-1} = \chi_\beta \oplus \chi_\alpha,$$

Reducing modulo λ, we see that

$$\{\overline{\alpha}_v, \overline{\beta}_v\} = \{\sharp k(v)^{1/2}\iota^{-1}(\chi_1(\varpi_v)), \sharp k(v)^{1/2}\iota^{-1}(\chi_2(\varpi_v))\}.$$

As a consequence, we see that $\chi_1\chi_2^{-1} \neq |\cdot|^{\pm 1}$. (Indeed, if $\chi_1\chi_2^{-1} = |\cdot|^{\pm 1}$, then $\overline{\alpha}_v\overline{\beta}_v^{-1} \equiv \sharp k(v)^{\pm 1} \equiv 1 \pmod{\lambda}$. This contradicts to our assumption.) Therefore, we have $\pi_v = I(\chi_1, \chi_2) \cong I(\chi_2, \chi_1)$. Without loss of generality, we assume that $\overline{\chi}_1(\varpi_v) = \overline{\beta}_v$ and $\overline{\chi}_2(\varpi_v) = \overline{\alpha}_v$.

We see that $S_Q \otimes_{\mathcal{O},\iota} \mathbb{C} = \oplus_\pi(\otimes_{v\in Q} X_v)$, where X_v is the one-dimensional space where U_{ϖ_v} acts via a lift of $\overline{\alpha}_v$. Since this space is spanned by ϕ_w, we see that Δ_v acts on S_Q via $\chi_2 = \chi_\alpha \circ \mathrm{Art}$. We then have proved the following result.

Lemma 5.2. *The two homomorphisms $\Delta_Q \to \mathrm{End}(S_Q)$ are equal.*

Let $U_{Q,0} := \prod_{v\notin Q} U_{Q,v} \prod_{v\in Q} U_0(v)$. Then U_Q is a normal subgroup of $U_{Q,0}$ and $U_{Q,0}/U_Q = \Delta_Q$. The following lemma is immediate from Corollary 3.17.

Lemma 5.3. *S_Q is finite free over $\mathcal{O}[\Delta_Q]$.*

Fix a place $v \in Q$. Since $\overline{\alpha}_v \neq \overline{\beta}_v$, by Hensel's lemma, we may write

$$\mathrm{Char}\,\rho_\emptyset^{\mathrm{mod}}(\mathrm{Frob}_v) = (X - A_v)(X - B_v)$$

for some $A_v,\ B_v \in \mathbb{T}_\emptyset$ with $A_v \equiv \overline{\alpha}_v,\ B_v \equiv \overline{\beta}_v \pmod{\mathfrak{m}_\emptyset}$.

Lemma 5.4. *The map*

$$\prod_{v\in Q}(U_{\varpi_v} - B_v) : S_\emptyset \to S(U_{Q,0}, \mathcal{O})_{\mathfrak{m}_Q}$$

is an isomorphism. Here we view the source and the target as submodules of $S(U_{Q,0}, \mathcal{O})_{\mathfrak{m}_\emptyset}$.

Proof. We claim that it is enough to prove that the map is an isomorphism after tensoring with L, and an injection after tensoring with $\mathbb{F}$. To see this, write $X := S_\emptyset$, $Y := S(U_{Q,0}, \mathcal{O})_{\mathfrak{m}_Q}$, and write Q for the cokernel of the map $X \to Y$. As X and Y are finite free $\mathcal{O}$-modules, if the map $X \otimes L \to Y \otimes L$ is injective, then so is the map $X \to Y$ and we have a short exact sequence

$$0 \to X \to Y \to Q \to 0.$$

Tensoring with L, we have $Q \otimes L = 0$. Tensoring with $\mathbb{F}$, we obtain an exact sequence

$$0 \to Q[\lambda] \to X \otimes \mathbb{F} \to Y \otimes \mathbb{F} \to Q \otimes \mathbb{F},$$

so we have $Q[\lambda] = 0$. Thus $Q = 0$ as required.

To check that we have an isomorphism after tensoring with L, it suffices to check that the induced map

$$\prod_{v\in Q}(U_{\varpi_v} - B_v) : S_\emptyset \otimes_{\mathcal{O},\iota} \mathbb{C} \to S(U_{Q,0}, \mathcal{O})_{\mathfrak{m}_Q} \otimes_{\mathcal{O},\iota} \mathbb{C}$$

is an isomorphism. This is easily checked. Indeed, we have

$$S_\emptyset \otimes \mathbb{C} \cong \oplus_\pi(\otimes_{v\in Q} I(\chi_{1,v}, \chi_{2,v}))^{\mathrm{GL}_2(\mathcal{O}_{F_v})}, \quad S(U_{Q,0}, \mathcal{O})_{\mathfrak{m}_Q} \otimes \mathbb{C} \cong \oplus_\pi(\otimes_{v\in Q} M_v).$$

Here M_v is the subspace of $I(\chi_{1,v},\chi_{2,v}))^{U_0(v)}$ on which U_{ϖ_v} acts via a lift of $\overline{\alpha}_v$, which is spanned by ϕ_w. On the other hand, we have computed that $I(\chi_{1,v},\chi_{2,v}))^{\mathrm{GL}_2(\mathcal{O}_{F_v})} = \mathbb{C}\phi_0$ (cf. Remark 3.7). Since the natural map $I(\chi_{1,v},\chi_{2,v}))^{\mathrm{GL}_2(\mathcal{O}_{F_v})} \to I(\chi_{1,v},\chi_{2,v}))^{U_0(v)}$ sends ϕ_0 to $\phi_1+\phi_w$, it suffices to check that $(U_{\varpi_v}-B_v)(\phi_1+\phi_w)$ is nontrivial. This follows from equation (5.2).

It remains to check the injectivity after tensoring with $\mathbb{F}$. The kernel of the map would be a nonzero finite module for the Artinian local ring $\mathbb{T}_\emptyset/\lambda$, and would thus have nonzero $\mathfrak{m}_\emptyset$-torsion, so it suffices to prove that the induced map

$$\prod_{v\in Q}(U_{\varpi_v}-B_v) : (S_\emptyset\otimes\mathbb{F})[\mathfrak{m}_\emptyset] \to S(U_{Q,0},\mathcal{O})_{\mathfrak{m}_Q}\otimes\mathbb{F}$$

is an injection. By induction on the number of elements in Q, it suffices to prove this in the case that $Q=\{v\}$. Suppose for the sake of contradiction that there is a nonzero $x\in(S_\emptyset\otimes\mathbb{F})[\mathfrak{m}_\emptyset]$ with $(U_{\varpi_v}-B_v)x=(U_{\varpi_v}-\overline{\beta}_v)x=0$. Since $x\in S_\emptyset\otimes\mathbb{F}$, we also have $T_vx=(\overline{\alpha}_v+\overline{\beta}_v)x$.

The Hecke operators are defined by double coset decomposition and we make them explicit. Note that there are elements g_i (cf. [84] and Remark 3.7) such that

$$U_v\begin{pmatrix}\varpi_v & 0\\ 0 & 1\end{pmatrix}U_v = \coprod_i g_iU_{Q,v},$$

$$\mathrm{GL}_2(\mathcal{O}_{F_v})\begin{pmatrix}\varpi_v & 0\\ 0 & 1\end{pmatrix}\mathrm{GL}_2(\mathcal{O}_{F_v}) = (\coprod_i g_i\,\mathrm{GL}_2(\mathcal{O}_{F_v}))\coprod\begin{pmatrix}1 & 0\\ 0 & \varpi_v\end{pmatrix}\mathrm{GL}_2(\mathcal{O}_{F_v}).$$

We have $\begin{pmatrix}1 & 0\\ 0 & \varpi_v\end{pmatrix}x = T_vx-U_{\varpi_v}x=\overline{\alpha}_vx$. Note that x is just a function $D^\times\backslash\mathrm{GL}_2(\mathbb{A}_F^\infty)\to\Lambda\otimes\mathbb{F}$, on which $\mathrm{GL}_2(\mathbb{A}_F^\infty)$ acts by right translation. We then have

$$\begin{pmatrix}\varpi_v & 0\\ 0 & \varpi_1\end{pmatrix}x = w\begin{pmatrix}1 & 0\\ 0 & \varpi_v\end{pmatrix}wx=\overline{\alpha}_vx,$$

and

$$U_{\varpi_v}x=\sum_{a\in k(v)}\begin{pmatrix}\varpi_v & \tilde{a}\\ 0 & 1\end{pmatrix}x=\sum_{a\in k(v)}\begin{pmatrix}1 & \tilde{a}\\ 0 & 2\end{pmatrix}\begin{pmatrix}\varpi_v & 0\\ 0 & 1\end{pmatrix}x=\sharp k(v)\overline{\alpha}_vx=\overline{\alpha}_vx.$$

Here $\tilde{a}\in\mathcal{O}_{F_v}$ is a lift of $a\in\mathbb{F}$. Then we obtain that $\overline{\alpha}_v=\overline{\beta}_v$, which is a contradiction. □

5.3. **The existence of auxiliary primes.** As promised, we prove the existence of auxiliary primes in the following.

Proposition 5.5. *Let* $r=\max\{\dim H^1(G_{F,T},(\mathrm{ad}^0\overline{\rho})(1)),\ 1+[F:\mathbb{Q}]-\sharp T\}$*. For each* $N\geq 1$*, there exists a set* Q_N *of primes of* F *such that*

- *$Q_N \cap T = \emptyset$.*
- *If $v \in Q_N$, then $\overline{\rho}(\mathrm{Frob}_v)$ has distinct eigenvalues $\overline{\alpha}_v \neq \overline{\beta}_v$.*
- *If $v \in Q_N$, then $\sharp k(v) \equiv 1 \pmod{p^N}$.*
- *$\sharp Q_N = r$.*
- *$R^{\square}_{Q_N}$ (resp. $R^{\square,\prime}_{Q_N}$ is topologically generated over R^{loc} (resp. $R^{\mathrm{loc},\prime}$) by $\sharp T - 1 - [F : \mathbb{Q}] + r$ elements.*

Proof. By what we have proved in [93], the last condition may be replaced by

$$H^1_{Q_N}(G_{F,T}, (\mathrm{ad}^0\,\overline{\rho})(1)) = (0).$$

Therefore, it suffices to show that for each $0 \neq [\phi] \in H^1(G_{F,T}, (\mathrm{ad}^0\,\overline{\rho})(1))$, there are infinitely many $v \notin T$ such that

(1) $\sharp k(v) \equiv 1 \pmod{p^N}$.
(2) $\overline{\rho}(\mathrm{Frob}_v)$ has distinct eigenvalues $\overline{\alpha}_v \neq \overline{\beta}_v$.
(3) $\mathrm{Res}[\phi] \in H^1(G_{k(v)}, (\mathrm{ad}^0\,\overline{\rho})(1))$ is nonzero.

Note that condition (1) is equivalent to v splitting completely in $F(\zeta_{p^N})$, condition (2) is equivalent to asking that $\mathrm{ad}\,\overline{\rho}(\mathrm{Frob}_v)$ has an eigenvalue not equal to 1.

Set $E = \overline{F}^{\mathrm{Ker}\,\mathrm{ad}\,\overline{\rho}}(\zeta_{p^N})$. We have assumed that $\mathrm{Im}\,\overline{\rho} \supseteq \mathrm{SL}_2(\mathbb{F}_p)$, from the classification of finite subgroups of $\mathrm{PGL}_2(\mathbb{F}_{p^s})$, this implies that $\mathrm{Im}\,\mathrm{ad}\,\overline{\rho} = \mathrm{PGL}_2(\mathbb{F}_{p^s})$ or $\mathrm{Im}\,\mathrm{ad}\,\overline{\rho} = \mathrm{PSL}_2(\mathbb{F}_{p^s})$ for some s. In particular $(\mathrm{Im}\,\mathrm{ad}\,\overline{\rho})^{\mathrm{ab}}$ is trivial or cyclic of order 2. Since $p \geq 5$ and p is unramified in F, we have $[F(\zeta_p) : F] \geq 4$. Therefore $\zeta_p \notin \overline{F}^{\mathrm{Ker}\,\mathrm{ad}\,\overline{\rho}}$.

The extension $E/\overline{F}^{\mathrm{Ker}\,\mathrm{ad}\,\overline{\rho}}$ is abelian. Let E_0 be the intermediate field such that $\mathrm{Gal}(E/E_0)$ has order prime to p and $\mathrm{Gal}(E_0/\overline{F}^{\mathrm{Ker}\,\mathrm{ad}\,\overline{\rho}})$ has p-power order. Write $\Gamma_1 = \mathrm{Gal}(E_0/F)$ and $\Gamma_2 = \mathrm{Gal}(E/E_0)$. We have the inflation-restriction exact sequence

$$0 \to H^1(\Gamma_1, (\mathrm{ad}^0\,\overline{\rho})(1)^{\Gamma_2}) \to H^1(\mathrm{Gal}(E/F), (\mathrm{ad}^0\,\overline{\rho})(1)) \to H^1(\Gamma_2, (\mathrm{ad}^0\,\overline{\rho})(1))^{\Gamma_1}.$$

Since E_0 contains $\overline{F}^{\mathrm{Ker}\,\mathrm{ad}\,\overline{\rho}}$, Γ_2 acts trivially on $\mathrm{ad}^0\,\overline{\rho}$. Therefore $(\mathrm{ad}^0\,\overline{\rho})(1)^{\Gamma_2} = 0$ since $\zeta_p \notin E_0$. Moreover, $H^1(\Gamma_2, (\mathrm{ad}^0\,\overline{\rho})(1)) = 0$ since Γ_2 has prime-to-p order. Hence the middle term $H^1(\mathrm{Gal}(E/F), (\mathrm{ad}^0\,\overline{\rho})(1)) = 0$.

Suppose that $\sharp k(v) \equiv 1 \pmod{p}$ and that $\overline{\rho}(\mathrm{Frob}_v) = \begin{pmatrix} \overline{\alpha}_v & 0 \\ 0 & \overline{\beta}_v \end{pmatrix}$. Then $\mathrm{ad}^0\,\overline{\rho}$ has the basis $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ of eigenvectors for Frob_v, with eigenvalues 1, $\overline{\alpha}_v/\overline{\beta}_v$, $\overline{\beta}_v/\overline{\alpha}_v$ respectively. Consequently, since $G_{k(v)}$ is a pro-cyclic group, we see that there is an isomorphism $H^1(G_{k(v)}, (\mathrm{ad}^0\,\overline{\rho})(1)) \cong \mathbb{F}$. This isomorphism is given by

$$[\phi] \mapsto \pi_v \circ \phi(\mathrm{Frob}_v) \circ i_v,$$

where i_v is the injection of $\mathbb{F}$ into the $\overline{\alpha}_v$-eigenspace of Frob_v, π_v is the Frob_v-equivariant projection onto that subspace.

Let $\sigma_0 \in \mathrm{Gal}(E/F)$ such that

(1) $\sigma_0(\zeta_{p^N}) = \zeta_{p^N}$,
(2) $\overline{\rho}(\sigma_0)$ has distinct eigenvalues $\overline{\alpha}$ and $\overline{\beta}$.

Indeed, such a σ_0 exists, because $\mathrm{Gal}(\overline{F}^{\mathrm{Ker}\,\overline{\rho}}/F(\zeta_{p^N}) \cap \overline{F}^{\mathrm{Ker}\,\overline{\rho}})$ contains $\mathrm{PSL}_2(\mathbb{F}_p)$ and then we may choose σ_0 so that its image in this group is an element whose adjoint has an eigenvalue other than 1.

Let $\tilde{E}/E$ be the extension cut out by all the $[\phi] \in H^1(G_{F,T}, (\mathrm{ad}^0\,\overline{\rho})(1))$. In order to complete the proof, it suffices to show that we can choose some $\sigma \in \mathrm{Gal}(\tilde{E}/F)$ with $\sigma|_E = \sigma_0$, and such that in the notation above, we have $\pi_{\sigma_0} \circ \phi(\sigma) \circ i_{\sigma_0} \neq 0$, because we can then choose v to have $\mathrm{Frob}_v = \sigma$ by the Cebotarev density theorem.

To this end, choose any $\tilde{\sigma}_0 \in \mathrm{Gal}(\tilde{E}/F)$ with $\tilde{\sigma}_0|_E = \sigma_0$. If $\tilde{\sigma}_0$ does not work, then we have $\pi_{\sigma_0} \circ \phi(\tilde{\sigma}_0) \circ i_{\sigma_0} \neq 0$. Take $\sigma = \sigma_1\tilde{\sigma}_0$ for some $\sigma_1 \in \mathrm{Gal}(\tilde{E}/E)$. Then

$$\phi(\sigma) = \phi(\sigma_1\tilde{\sigma}_0) = \phi(\sigma_1) + \sigma_1\phi(\tilde{\sigma}_0) = \phi(\sigma_1) + \phi(\tilde{\alpha}_0),$$

so $\pi_{\sigma_0} \circ \phi(\sigma) \circ i_{\sigma_0} = \pi_{\sigma_0} \circ \phi(\sigma_1) \circ i_{\sigma_0}$.

Note that $\phi(\mathrm{Gal}(\tilde{E}/E))$ is a $\mathrm{Gal}(E/F)$-invariant subset of $\mathrm{ad}^0\,\overline{\rho}$, which is an irreducible $\mathrm{Gal}(E/F)$-module as the image of $\overline{\rho}$ contains $\mathrm{SL}_2(\mathbb{F}_p)$. Thus the $\mathbb{F}$-span of $\phi(\mathrm{Gal}(\tilde{E}/E))$ is all of $\mathrm{ad}^0\,\overline{\rho}(1)$, from which it is immediate that we can choose σ_1 so that $\pi_{\sigma_0} \circ \phi(\sigma_1) \circ i_{\sigma_0} \neq 0$. This completes the proof. □

5.4. **Proof of Theorem 1.1.** Set $S_Q^\square := S_Q \otimes_{R_Q^{\mathrm{univ}}} R_Q^\square$. Then we have

$$S_Q^\square/\mathfrak{a}_Q = S(U_{Q,0}, \mathcal{O})_{\mathfrak{m}_Q} \xrightarrow{\sim} S_\emptyset,$$

compatibly with the isomorphism $R_Q^\square/\mathfrak{a}_Q \xrightarrow{\sim} R_\emptyset^{\mathrm{univ}}$. Moreover, $S_Q^\square$ is finite free over $\mathcal{J}[\Delta_Q]$.

We may and do choose a presentation

$$R^{\mathrm{loc}}[[x_1, \dots, x_{h_Q}]] \twoheadrightarrow R_Q^\square,$$

where $h_Q = \sharp T + \sharp Q - 1 - [F : \mathbb{Q}] + \dim_{\mathbb{F}} H_Q^1(G_{F,T}, (\mathrm{ad}^0\,\overline{\rho})(1))$, and

$$H_Q^1(G_{F,T}, (\mathrm{ad}^0\,\overline{\rho})(1)) = \mathrm{Ker}(H^1(G_{F,T}, (\mathrm{ad}^0\,\overline{\rho})(1)) \to \oplus_{v\in Q} H^1(G_{k(v)}, (\mathrm{ad}^0\,\overline{\rho})(1)).$$

Write $h := \sharp T - 1 - [F : \mathbb{Q}] + r$ and $R_\infty := R^{\mathrm{loc}}[[x_1, \dots, x_h]]$. For each set Q_N as above, choose a surjection $R_\infty \twoheadrightarrow R_{Q_N}^\square$. Let $\mathcal{J}_\infty := \mathcal{J}[[y_1, \dots, y_r]]$. Choose a surjection

$\mathcal{J}_\infty \twoheadrightarrow \mathcal{J}[\Delta_{Q_N}]$, given by writing $Q_N = \{v_1, \dots, v_r\}$ and mapping y_i to $(\gamma_i - 1)$, where γ_i is a generator of Δ_{v_i}. Choose a homomorphism $\mathcal{J}_\infty \to R_\infty$ so that the composites

$$\mathcal{J}_\infty \to R_\infty \twoheadrightarrow R^\square_{Q_N} \text{ and } \mathcal{J}_\infty \to \mathcal{J}[\Delta_{Q_N}] \to R^\square_{Q_N}$$

agree. Write $\mathfrak{a}_\infty := (\mathfrak{a}, y_1, \dots, y_r)$. Then

$$S^\square/\mathfrak{a}_\infty = S_\emptyset \text{ and } R^\square_{Q_N}/\mathfrak{a}_\infty = R^{\mathrm{univ}}_\emptyset.$$

Write $\mathfrak{b}_N := \mathrm{Ker}(\mathcal{J}_\infty \to \mathcal{J}[\Delta_{Q_N}])$, so that $S^\square_{Q_N}$ is finite free over $\mathcal{J}_\infty/\mathfrak{b}_N$. Since all the elements of Q_N are congruent to 1 modulo p^N, we see that

$$\mathfrak{b}_N \subset ((1+y_1)^{p^N} - 1, \dots, (1+y_r)^{p^N} - 1).$$

We may and do choose the same data for $R^{\mathrm{loc},\prime}$, in such a way that two sets of data are compatible modulo λ.

Choose open ideals $\mathfrak{c}_N \lhd \mathcal{J}_\infty$ such that

- $\mathfrak{c}_N \cap \mathcal{O} = (\lambda^N)$;
- $\mathfrak{c}_N \supset \mathfrak{b}_N$;
- $\mathfrak{c}_N \supset \mathfrak{c}_{N+1}$;
- $\cap_N \mathfrak{c}_N = (0)$.

For example we could take $\mathfrak{c}_N = ((1 + X_{v,i,j})^{p^N} - 1, \dots, (1 + y_i)^{p^N} - 1, \lambda^N)$. Note that since $\mathfrak{c}_N \supset \mathfrak{b}_N$, $S^\square_{Q_N}/\mathfrak{c}_N$ is finite free over $\mathcal{J}_\infty/\mathfrak{c}_N$. Choose open ideals $\mathfrak{d}_N \lhd R^{\mathrm{univ}}_\emptyset$ such that

- $\mathfrak{d}_N \subset \mathrm{Ker}(R^{\mathrm{univ}}_\emptyset \to \mathrm{End}(S_\emptyset/\lambda^N))$;
- $\mathfrak{d}_N \supset \mathfrak{d}_{N+1}$;
- $\cap_N \mathfrak{d}_N = (0)$.

If $M \geq N$, write $S_{M,N} := S^\square_{Q_M}/\mathfrak{c}_N$, so that $S_{M,N}$ is finite free over $\mathcal{J}_\infty/\mathfrak{c}_N$ of rank equal to the $\mathcal{O}$-rank of $S_\emptyset$. Indeed, this follows from the isomorphism $S_{M,N}/\mathfrak{a}_\infty \xrightarrow{\sim} S_\emptyset/\lambda^N$. Then we have a commutative diagram

$$\begin{array}{ccccc} \mathcal{J}_\infty & \longrightarrow & R_\infty & \longrightarrow\!\!\!\!\!\rightarrow & R^{\mathrm{univ}}_\emptyset/\mathfrak{d}_N \\ & & \downarrow & & \downarrow \\ & & S_{M,N} & \longrightarrow\!\!\!\!\!\rightarrow & S_\emptyset/\mathfrak{d}_N. \end{array}$$

Here the dotted arrows mean the module structure, the objects $S_{M,N}$ and $S_\emptyset/\mathfrak{d}_N$ have finite cardinality. Therefore, there is an infinite subsequence of pairs (M_i, N_i) such that

$M_{i+1} > M_i$, $N_{i+1} > N_i$, $M_i \geq N_i$, and the induced diagram

$$\begin{array}{ccccc} \mathcal{J}_\infty & \longrightarrow & R_\infty & \twoheadrightarrow & R_\emptyset^{\mathrm{univ}}/\mathfrak{d}_{N_i} \\ & & \downarrow & & \downarrow \\ & & S_{M_{i+1},N_{i+1}}/\mathfrak{c}_{N_i} & \twoheadrightarrow & S_\emptyset/\mathfrak{d}_{N_i}. \end{array}$$

is isomorphic to the diagram for (M_i, N_i). Then we could take the projective limit over this subsequence, to obtain a commutative diagram

$$\begin{array}{ccccc} \mathcal{J}_\infty & \longrightarrow & R_\infty & \twoheadrightarrow & R_\emptyset^{\mathrm{univ}} \\ & & \downarrow & & \downarrow \\ & & S_\infty & \twoheadrightarrow & S_\emptyset. \end{array}$$

Here S_∞ is finite free over $\mathcal{J}_\infty$. Furthermore, we can carry out the same construction in the $'$ world, compatibly with this picture modulo λ.

Now we could deduce our main result by purely commutative algebra arguments. We have

$$\dim R_\infty = \dim R'_\infty = \dim \mathcal{J}_\infty = 4\sharp T + r,$$

and since S_∞ and S'_∞ are finite free over the power series ring $\mathcal{J}_\infty$, we have

$$\mathrm{depth}_{\mathcal{J}_\infty}(S_\infty) = \mathrm{depth}_{\mathcal{J}_\infty}(S'_\infty) = 4\sharp T + r.$$

Since the action of $\mathcal{J}_\infty$ on S_∞ factors through R_∞, we see that

$$\mathrm{depth}_{R_\infty}(S_\infty) \geq 4\sharp T + r,$$

and similarly

$$\mathrm{depth}_{R'_\infty}(S'_\infty) \geq 4\sharp T + r.$$

If $\mathcal{P} \lhd R'_\infty$ is a minimal prime in the support of S'_∞, then we have

$$4\sharp T + r = \dim R'_\infty \geq \dim R'_\infty/\mathcal{P} \geq \mathrm{depth}_{R'_\infty} S'_\infty \geq 4\sharp T + r,$$

hence equality holds throughout and $\mathcal{P}$ is a minimal prime of R'_∞. But R'_∞ has a unique minimal prime, so in fact $\mathrm{Supp}_{R'_\infty}(S'_\infty) = \mathrm{Spec}\, R'_\infty$.

By the same argument, $\mathrm{Supp}_{R_\infty}(S_\infty)$ is a union of irreducible components of $\mathrm{Spec}\, R_\infty$. We show that it is all of $\mathrm{Spec}\, R_\infty$ by reducing modulo λ and comparing with the situation for S'_∞.

Indeed, since $\mathrm{Supp}_{R'_\infty}(S'_\infty) = \mathrm{Spec}\, R'_\infty$, we certainly have $\mathrm{Supp}_{R'_\infty/\lambda}(S'_\infty/\lambda) = \mathrm{Spec}\, R'_\infty/\lambda$. By the compatibility between the two pictures, this means that

$$\mathrm{Supp}_{R_\infty/\lambda}(S_\infty/\lambda) = \mathrm{Spec}\, R_\infty/\lambda.$$

Thus $\operatorname{Supp}_{R_\infty/\lambda}(S_\infty/\lambda)$ is a union of irreducible components of $\operatorname{Spec} R_\infty$, which contains the entirety of $\operatorname{Spec} R_\infty/\lambda$. Since the irreducible components of $\operatorname{Spec} R_\infty/\lambda$ are in bijection with the irreducible components of $\operatorname{Spec} R_\infty$, this implies that

$$\operatorname{Supp}_{R_\infty}(S_\infty) = \operatorname{Spec} R_\infty.$$

Therefore $\operatorname{Supp}_{R_\infty/\mathfrak{a}_\infty}(S_\infty/\mathfrak{a}_\infty) = \operatorname{Spec} R_\infty/\mathfrak{a}_\infty$, i.e. $\operatorname{Supp}_{R_\emptyset^{\text{univ}}} S_\emptyset = \operatorname{Spec} R_\emptyset^{\text{univ}}$, which proves the modularity by Lemma 5.1.

6. Ultrapatching

The patching procedure in Section 4 works as a program: input rings and modules satisfying certain conditions, we obtain objects with "good" properties. In this section we summarize the commutative algebra behind the patching method and explain the *ultrapatching* construction introduced by Scholze in [68], see also [57, 58].

Let X be a set. An filter on X is a consistent choice of which subsets of X are "large".

Definition 6.1. A *filter* on X is $\mathcal{F} \subset \mathcal{P}(X)$ such that

(1) $X \in \mathcal{F}$ (the whole set is large);
(2) $\emptyset \notin \mathcal{F}$ (the empty set is not large);
(3) If $A \in \mathcal{F}$ and $A \subset B$, then $B \in \mathcal{F}$ (any set containing a large set is large);
(4) If $A,\ B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$ (large sets have large intersection).

A filter is *principal* if $\mathcal{F} = \{A \subset X \mid x \in A\}$ for some $x \in X$.

A filter is an *ultrafilter* if for any $A \subset X$, $A \in \mathcal{F}$ or $X - A \in \mathcal{F}$.

It is well known that for $X = \mathbb{N}$, nonprincipal unltrafilters $\mathcal{F}$ exist provided we assume the axiom of choice. From now on, we fix a nonprincipal ultrafilter $\mathfrak{F}$ on $\mathbb{N}$. For convenience, we say that a property $\mathcal{P}(n)$ holds for $\mathfrak{F}$-many i if there is some $I \in \mathfrak{F}$ such that $\mathcal{P}(i)$ holds for all $i \in I$.

Definition 6.2. For any sequence of sets $\mathscr{A} = \{A_n\}_{n\geq 1}$, their *ultraproduct* is the quotient

$$\mathcal{U}(\mathscr{A}) := (\prod_{n=1}^{\infty} A_n)/\sim,$$

where the equivalence relation $\sim$ is defined by $(a_n)_n \sim (a'_n)_n$ if $a_i = a'_i$ for $\mathfrak{F}$-many i.

Remark 6.3. If the A_n's are sets with an algebraic structure (e.g. groups, rings, R-modules, R-algebras, etc.) then $\mathcal{U}(\mathscr{A})$ naturally inherits the same structure.

If each A_n is a finite set and the cardinalities of the A_n's are bounded, then $\mathcal{U}(\mathscr{A})$ is also a finite set and there are bijections $\mathcal{U}(\mathscr{A}) \xrightarrow{\sim} A_i$ for $\mathfrak{F}$-many i. Moreover, if the

A_n's are sets with an algebraic structure, such that there are only finitely many distinct isomorphism classes appearing in $\{A_n\}_{n\geq 1}$ (which happens automatically if the structure is defined by finitely many operations, e.g. groups, rings, R-modules and R-algebras over a finite ring R), then these bijections may be taken to be isomorphisms. Indeed, by our assumption, there is some A such that $A \cong A_i$ for $\mathfrak{F}$-many i and then $\mathcal{U}(\mathscr{A})$ is isomorphic to the unltraproduct $\mathcal{U}(\{A\}_{n\geq 1})$, which is isomorphic to A if A is a finite set.

In the case when each A_n is a module over a finite local ring R, there is a simple algebraic description of $\mathcal{U}(\mathscr{A})$. Specifically, the ring $\mathcal{R} = \prod_{n=1}^{\infty} R$ contains a unique maximal ideal $\mathfrak{I}$ for which $\mathcal{R}_{\mathfrak{I}} \cong R$ and $(\prod_{n=1}^{\infty} A_n)_{\mathfrak{I}} \cong \mathcal{U}(\mathscr{A})$ as R-modules. This shows that $\mathcal{U}(-)$ is a particularly well-bahaved functor in this special situation. In particular, it is exact.

In the following, fix a power series ring $S_\infty = \mathcal{O}[[z_1, \dots, z_d]]$ and consider the ideal $\mathfrak{n} = (z_1, \dots, z_d)$. Fix a sequence of ideals $\mathcal{I}_n \subset S_\infty$ such that for any open ideal $\mathfrak{a} \subset S_\infty$ we have $\mathcal{I}_n \subset \mathfrak{a}$ for all but finitely many n. Define

$$\overline{S}_\infty = S_\infty/(\varpi) = \mathbb{F}[[z_1, \dots, z_d]] \text{ and } \overline{\mathcal{I}}_n = (\mathcal{I}_n + (\varpi))/(\varpi) \subset \overline{S}_\infty.$$

For any finitely generated S_∞-module M, the *S_∞-rank* of M, denoted by $\operatorname{rank}_{S_\infty} M$, is defined to be the cardinality of a minimal generating set for M as an S_∞-module.

Definition 6.4. Let $\mathscr{M} = \{M_n\}_{n\geq 1}$ be a sequence of finitely generated S_∞-modules with $\mathcal{I}_n \subset \operatorname{Ann}_{S_\infty} M_n$ for all but finitely many n.

- We say that $\mathscr{M}$ is a *weak patching system* if the S_∞-ranks of the M_n's are uniformly bounded. If we further have $\varpi M_n = 0$ for all but finitely many n, we say that $\mathscr{M}$ is a *residual weak patching system*.
- We say that $\mathscr{M}$ is a *patching system* if it is a weak patching system and $\operatorname{Ann}_{S_\infty} M_n = \mathcal{I}_n$ for all but finitely many n.
- We say that $\mathscr{M}$ is a *residual patching system* if it is a patching system and $\operatorname{Ann}_{\overline{S}_\infty} M_n = \overline{\mathcal{I}}_n$ for all but finitely many n.
- We say that $\mathscr{M}$ is MCM (resp. MCM residual) if $\mathscr{M}$ is a patching system (resp. residual patching system) and M_n is free over $S_\infty/\mathcal{I}_n$ (resp. $\overline{S}_\infty/\overline{\mathcal{I}}_n$) for all but finitely many n.

Furthermore, assume that $\mathscr{R} = \{R_n\}_{n\geq 1}$ is a sequence of finite type local S_∞-algebras.

- We say that $\mathscr{R} = \{R_n\}_{n\geq 1}$ is a (*weak*, *residual*) *patching algebra*, if it is a (weak, residual) patching system.

- If M_n is an R_n-module (viewed as an S_∞-module via the S_∞-algebra structure on R_n) for all n, we say that $\mathscr{M} = \{M_n\}_{n\geq 1}$ is a (*weak*, *residual*) *patching* $\mathscr{R}$*-module* if it is a (weak, residual) patching system.

Let $\mathfrak{w}\mathfrak{P}$ be the category of weak patching systems, with the obvious notion of morphism. It is naturally an abelian category.

Definition 6.5. Let $\mathscr{M}$ be a weak patching system. The *patched module* of $\mathscr{M}$ is the S_∞ module

$$\mathscr{P}(\mathscr{M}) = \varprojlim_{\mathfrak{a}} \mathcal{U}(\mathscr{M}/\mathfrak{a}),$$

where the inverse limit is taken over all open ideals of S_∞. We may treat $\mathscr{P}$ as a functor from $\mathfrak{w}\mathfrak{P}$ to the category of S_∞-modules.

Remark 6.6. If $\mathscr{R}$ is a weak patching algebra and $\mathscr{M}$ is a weak patching $\mathscr{R}$-module, then $\mathscr{P}(\mathscr{R})$ inherits a natural S_∞-algebra structure, and $\mathscr{P}(\mathscr{M})$ inherits a natural $\mathscr{P}(\mathscr{R})$-module structure.

In the above definition, the ultraproduct essentially plays the role of the pigeonhole principal in the classical Taylor-Wiles-Kisin construction (cf. Proposition 4.1), with the simplification that it is not necessary to explicitly define a patching datum before making the construction. Indeed, if one were to define patching data for the $M_n/\mathfrak{a}$'s, then the machinery of ultraproducts would ensure that the patching data for $\mathcal{U}(\mathscr{M}/\mathfrak{a})$ would agree with that of $M_n/\mathfrak{a}$ for infinitely many n. Proposition 4.1 then can be rephrased as follows.

Proposition 6.7. *let $\mathscr{R}$ be a weak patching algebra, $\mathscr{M}$ be an MCM patching $\mathscr{R}$-module. Then*

1. *$\mathscr{P}(\mathscr{R})$ is a finite type S_∞-algebra, and $\mathscr{P}(\mathscr{M})$ is a finitely generated free S_∞-module.*
2. *The structure map $S_\infty \to \mathscr{P}(\mathscr{R})$ is injective, and thus $\dim \mathscr{P}(\mathscr{R}) = \dim S_\infty$.*
3. *The module $\mathscr{P}(\mathscr{M})$ is maximal Cohen-Macaulay over $\mathscr{P}(\mathscr{R})$, and $(\varpi, z_1, \ldots, z_d)$ is a regular sequence for $\mathscr{P}(\mathscr{M})$.*

Proposition 6.8. *let $\mathscr{R}$ be a weak patching algebra, $\overline{\mathscr{M}}$ be an MCM residual patching $\mathscr{R}$-module. Then*

1. *$\mathscr{P}(\mathscr{R})/(\varpi)$ is a finite type $\overline{S}_\infty$-algebra, and $\mathscr{P}(\overline{\mathscr{M}})$ is a finitely generated free $\overline{S}_\infty$-module.*
2. *The structure map $\overline{S}_\infty \to \mathscr{P}(\mathscr{R})/(\varpi)$ is injective, and thus $\dim \mathscr{P}(\mathscr{R})/(\varpi) = \dim \overline{S}_\infty$.*

(3) *The module* $\mathscr{P}(\overline{\mathscr{M}})$ *is maximal Cohen-Macaulay over* $\mathscr{P}(\mathscr{R})/(\varpi)$*, and* $(z_1, \ldots, z_d)$ *is a regular sequence for* $\mathscr{P}(\overline{\mathscr{M}})$*.*

Proposition 6.9. *Let* $\mathfrak{n} = (z_1, \ldots, z_d) \subset S_\infty$ *as above. Let* R_0 *be a finite type* $\mathcal{O}$*-algebra,* M_0 *be a finitely generated* R_0*-module. If, for each* $n \geq 1$*, there are isomorphisms* $R_n/\mathfrak{n} \cong R_0$ *of* $\mathcal{O}$*-algebras and* $M_n/\mathfrak{n} \cong M_0$ *of* $R_n/\mathfrak{n} \cong R_0$*-modules, then we have* $\mathscr{P}(\mathscr{R})/\mathfrak{n} \cong R_0$ *as* $\mathcal{O}$*-algebras and* $\mathscr{P}(\mathscr{M})/\mathfrak{n} \cong M_0$ *as* $\mathscr{P}(\mathscr{R})/\mathfrak{n} \cong R_0$*-modules.*

From the set up of Proposition 6.7, there is very little we can conclude about the ring $\mathscr{P}(\mathscr{R})$. However in practice one generally takes the rings R_n to be quotients of a fixed ring R_∞ of the same dimension as S_∞. Thus we define a *cover* of a weak patching algebra $\mathscr{R} = \{R_n\}_{n\geq 1}$ to be a pair $(R_\infty, \{\varphi_n\}_{n\geq 1})$, where R_∞ is a complete, topologically finitely generated $\mathcal{O}$-algebra of Krull dimension $\dim S_\infty$, and $\varphi_n : R_\infty \to R_n$ is a surjective $\mathcal{O}$-algebra homomorphism for each n. It is straightforward to show the following (cf. [57]).

Proposition 6.10. *If* $(R_\infty, \{\varphi_n\}_{n\geq 1})$ *is a cover of a weak patching algebra* $\mathscr{R}$*, then the* φ_n*'s induce a natural continuous surjection* $\varphi_\infty : R_\infty \to \mathscr{P}(\mathscr{R})$*.*

Using the fact [91, Lemma 0AAD] (which says that if $f : A \twoheadrightarrow B$ is a surjection of noetherian local rings, then a B-module M is Cohen-Macaulay as an A-module if and only if it is Cohen-Macaulay as a B-module), combining Propositions 6.7, 6.8, and 6.10, we have the following result.

Corollary 6.11. *Let* $\mathscr{R}$ *be a weak patching algebra and let* $(R_\infty, \{\varphi_n\}_{n\geq 1})$ *be a cover of* $\mathscr{R}$*. If* $\mathscr{M}$ *is an MCM patching* $\mathscr{R}$*-module, then* $\mathscr{P}(\mathscr{M})$ *is a maximal Cohen-Macaulay* R_∞*-module. If* $\overline{\mathscr{M}}$ *is an MCM residual patching* $\mathscr{R}$*-module, then* $\mathscr{P}(\overline{\mathscr{M}})$ *is a maximal Cohen-Macaulay* $R_\infty/(\varpi)$*-module.*

If $\mathscr{P}$ were an exact functor, then we could patch objects with respect to certain filtration. However, this is not true in general.

Example 6.12. Assume that $S_\infty/\mathcal{I}_n$ is ϖ-torsion free for all n and let $\mathscr{M} = \{S_\infty/\mathcal{I}_n\}_{n\geq 1}$. Define $\varphi = \{\varphi_n\}_{n\geq 1} : \mathscr{M} \to \mathscr{M}$ by $\varphi_n(x) = \varpi^n x$. Then $\phi : \mathscr{M} \to \mathscr{M}$ is injective, $\mathscr{P}(\mathscr{M}) = S_\infty$, and $\mathscr{P}(\varphi) : S_\infty \to S_\infty$ is the zero map.

Nevertheless we have the following weaker statement.

Lemma 6.13. *Then functor* $\mathcal{P}(-)$ *is right exact. If*

$$0 \to \mathscr{A} \to \mathscr{B} \to \mathscr{C} \to 0$$

is an exact sequence of weak patching systems, then

$$0 \to \mathscr{P}(\mathscr{A}) \to \mathscr{P}(\mathscr{B}) \to \mathscr{P}(\mathscr{C}) \to 0$$

is exact, provided that either:

- $\mathscr{C}$ *is MCM, or*
- $\mathscr{A}$, $\mathscr{B}$, *and* $\mathscr{C}$ *are all residual weak patching systems, and* $\mathscr{C}$ *is MCM residual.*

Proof. Let $\underline{\mathrm{Ab}}$ be the category of abelian groups. For any countable directed set I, let $\underline{\mathrm{finAb}}^I$ be the category of inverse systems of finite abelian groups indexed by I.

Let $(A_i, f_{ji} : A_j \to A_i)$ be an object in $\underline{\mathrm{finAb}}^I$. Since A_i is finite and $\{\mathrm{Im}\, f_{ji}\}_{j \geq i}$ is a decreasing sequence of subgroups of A_i, there is a $j \geq i$ for which $\mathrm{Im}\, f_{ki} = \mathrm{Im}\, f_{ji}$ for all $k \geq j$. I.e. the inverse system $(A_i, f_{ji} : A_j \to A_i)$ satisfies the Mittag-Leffler condition.

Assume that we have an exact sequence of weak patching systems

$$0 \to \mathscr{A} \to \mathscr{B} \to \mathscr{C} \to 0.$$

For any $\mathfrak{a} \subset S_\infty$, the sequence

$$\mathscr{A}/\mathfrak{a} \to \mathscr{B}/\mathfrak{a} \to \mathscr{C}/\mathfrak{a} \to 0$$

is exact. By the exactness of $\mathcal{U}(-)$, we obtain the exact sequence

$$\mathcal{U}(\mathscr{A}/\mathfrak{a}) \to \mathcal{U}(\mathscr{B}/\mathfrak{a}) \to \mathcal{U}(\mathscr{C}/\mathfrak{a}) \to 0,$$

hence an exact sequence of inverse systems

$$(\mathcal{U}(\mathscr{A}/\mathfrak{a}))_\mathfrak{a} \to (\mathcal{U}(\mathscr{B}/\mathfrak{a}))_\mathfrak{a} \to (\mathcal{U}(\mathscr{C}/\mathfrak{a}))_\mathfrak{a} \to 0.$$

Note that there are only countably many open ideals of S_∞, and $\mathcal{U}(\mathscr{A}/\mathfrak{a})$, $\mathcal{U}(\mathscr{B}/\mathfrak{a})$, and $\mathcal{U}(\mathscr{C}/\mathfrak{a})$ are all finite. Taking inverse limits preserves exactness, i.e. the sequence

$$\mathscr{P}(\mathscr{A}) \to \mathscr{P}(\mathscr{B}) \to \mathscr{P}(\mathscr{C}) \to 0$$

is exact. The functor $\mathcal{P}(-)$ is right exact.

Now assume that one of the further conditions of the lemma holds. Write $\mathscr{A} = \{A_n\}_{n\geq 1}$, $\mathscr{B} = \{B_n\}_{n\geq 1}$, $\mathscr{C} = \{C_n\}_{n\geq 1}$. Let $I_n = \mathrm{Ann}_{S_\infty} C_n$ (so that either $I_n = \mathcal{I}_n$ or $\overline{\mathcal{I}}_n$ for all $n \geq 1$), we have, for all $n \gg 0$, an exact sequence of S_∞-modules

$$0 \to A_n \to B_n \to C_n \to 0,$$

and C_n is a free S_∞/I_n-module. It follows that

$$\mathrm{Tor}_1^{S_\infty/I_n}(C_n, S_\infty/\mathfrak{a}) = 0$$

for all $\mathfrak{a} \subset S_\infty$. Therefore,

$$0 \to A_n/\mathfrak{a} \to B_n/\mathfrak{a} \to C_n/\mathfrak{a} \to 0$$

is exact for all $n \gg 0$. Then it is easy to see that the sequence

$$0 \to \mathscr{P}(\mathscr{A}) \to \mathscr{P}(\mathscr{B}) \to \mathscr{P}(\mathscr{C}) \to 0$$

is exact. □

We have the following useful consequence.

Corollary 6.14. *Let $\mathscr{V}$ be a residual weak patching system with a filtration*

$$0 = \mathscr{V}^0 \subset \mathscr{V}^1 \subset \cdots \subset \mathscr{V}^r = \mathscr{V}$$

by residual weak patching systems $\mathscr{V}^k$. For $k = 1, \ldots, r$, let $\mathscr{M}^k = \mathscr{V}^k/\mathscr{V}^{k-1}$. Assume that the $\mathscr{M}^k$'s are all MCM residual. Then $\mathscr{P}(\mathscr{V})$ has a filtration

$$0 = \mathscr{P}(\mathscr{V}^0) \subset \mathscr{P}(\mathscr{V}^1) \subset \cdots \subset \mathscr{P}(\mathscr{V}^r) = \mathscr{P}(\mathscr{V})$$

with $\mathscr{P}(\mathscr{M}^k) = \mathscr{P}(\mathscr{V}^k)/\mathscr{P}(\mathscr{V}^{k-1})$ for all $k = 1, \ldots, r$.

References

[1] J. Arthur, L. Clozel: *Simple Algebras, Base Change, and the Advanced Theory of Trace Formula.* Annals of Math. Studies, No. 120, Princeton University Press, 1989.

[2] M. W. Baldoni: *General representation theory of real reductive Lie groups.* In Representation Theory and Automorphic Forms, Proceedings of Symposia in Pure Mathematics, vol. 61, 1997.

[3] A. Borel: *Some finiteness properties of adele groups over number fields.* Publ. Math. I.H.E.S. 16 (1963), 5-30.

[4] A. Borel; Harish-Chandra: *Arithmetic subgroups of algebraic groups.* Annals of Math. 75 (1962), 485-535.

[5] D. Bump: *Automorphic forms and represnetations.* Cambridge University Press 1998. ISBN: 0-521-55098-X

[6] T. Barnet-Lamb; T. Gee; D. Geraghty; R. Taylor: *Potential automorphy and change of weight.* Annals of Math 179 (2014), 501-609.

[7] T. Barnet-Lamb; D. Geraghty; M. Harris; R. Taylor: *A family of Calabi-Yau varieties and potential automorphy 2.* P.R.I.M.S. 47 (2011), 29-98.

[8] A. Borel; H. Jacquet: Automorphic Forms, Representations, and L-Functions, Proceedings of Symposia in Pure Mathematics, vol. 33, part 1, 1979.

[9] C. Breuil: *Groupes p-divisibles, groupes finis et modules filtrés.* Ann. of Math. (2) 152 (2000), 489-549.

[10] C. Breuil; B. Conrad; F. Diamond; R. Taylor: *On the modularity of elliptic curves over $\mathbb{Q}$: wild 3-adic exercises.* J. Amer. Math. Soc. 14 (2001), 843-939.

[11] C. Breuil; F. Diamond: *Hilbert modular forms modulo p and values of extensions between Galois characters.* Ann. Sci. Éc. Norm. Supér. (4) 47 (2014), no. 5, 905-974.

[12] C. Breuil; A. Mézard: *Multiplicités modulaires et repreésentations de* $\mathrm{GL}_2(\mathbb{Z}_p)$ *et de* $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ *en* $l = p$. Duke Math. J. 115 (2002), 205-310.

[13] W. Bruns, J. Herzog: *Cohen-Macaulary rings.* Cambridge Studies in Adv. Math. 39, Cambridge University Press, Cambridge (1986).

[14] C. J. Bushnell, G. Henniart: *The local Langlands conjecture for* GL(2). ISBN 978-3-540-31511-7.

[15] K. Buzzard: *Potential modularity–a survey.* In *Non-abelian fundamental groups and Iwasawa theory*, 188–211, London Math. Soc. Lecture Note Ser., 393, Cambridge Univ. Press, Cambridge, 2012.

[16] K. Buzzard; F. Diamond; F. Jarvis: *On Serre's conjecture for mod* ℓ *Galois representations over totally real fields*. Duke Math. J. Vol. 155, no. 1, 2010, 105-161.

[17] H. Carayol: *Sur les représentations l-adiques attachées aux formes modulaires de Hilbert.* C. R. Acad. Sci. Paris Sér. I Math. 296 (1983), no. 15, 629–632.

[18] W. Casselman: *On some results of Atkin and Lehner.* Math. Ann. 201 (1973) 301-314.

[19] L. Clozel; M. Harris; R. Taylor: *Automorphy for some l-adic lifts of automorphic mod l representations.* Pub. Math. IHES 108 (2008), 1-181.

[20] G. Cornell; J. H. Silverman; G. Stevens: *Modular forms and Fermat's Last Theorem.* Springer 1997.

[21] B. Conrad; F. Diamond; R. Taylor: *Modularity of certain potentially Barsotti-Tate Galois representations.* J. Amer. Math. Soc. 12 (1999), no. 2, 521–567.

[22] H. Darmon; F. Diamond; R. Taylor: *Fermat's last theorem.* In *Elliptic curves, modular forms and Fermat's last theorem* (Hong Kong, 1993), 2-140, Int. Press, Cambridge, MA, 1997.

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

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

[25] F. Diamond; J. Shurman: *A first course in modular forms.* Graduate Texts in Mathematics, 228. Springer-Verlag, New York, 2005. xvi+436 pp. ISBN: 0-387-23229-X.

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

[27] M. Emerton: *Local-global compatiblity in the p-adic Langlands program fro* GL_2. `http://www.math.uchicago.edu/~emerton/pdffiles/lg.pdf`

[28] D. Flath: *Decomposition of representations into tensor products.* In Automorphic Forms, Representations, and L-functions, Proceedings of Symposia in Pure Mathematics, vol. 33, part 1, 1979.

[29] J-M. Fontaine; G. Laffaille: *Construction de représentations p-adiques.* Ann. Sci. école Norm. Sup. (4) 15 (1982), no. 4, 547-608 (1983).

[30] N. Freitas; B. V. Le Hung; S. Siksek: *Elliptic curves over real quadratic fields are modular.* Inventiones Mathematicae, 201 (1): 159-206.

[31] G. Frey: *Links between solutions of* $A - B = C$ *and elliptic curves.* In *Number Theory, proceedings of the Journees arithmetiques, held in Ulm, 1987,* H. P. Schlichewei, E. Wirsing, editors. Lecture Notes in Mathematics 1380. Springer-Verlag, Berlin, New York, 1989.

[32] G. Frey: *Links between stabel elliptic curves and certain Diophantine equations.* Ann. Univ. Saraviensis, Ser. Math. 1 (1986), 1-40.

[33] K. Fujiwara: *Deformation rings and Hecke algebras in the totally real case.* `arxiv.org/abs/math/0602606`

[34] P. B. Garrett: *Holomorphic Hilbert modular forms.* Wadsworth and Brooks/Cole, Pacific Grove, CA, 1990.

[35] T. Gee: *Automorphic lifts of prescribed types.* Mathematische Annalen 350 (2011), 107-144.

[36] T. Gee: *Modularity lifting theorems.* Notes for Arizona Winter School 2013.

[37] T. Gee; M. Kisin: *The Breuil-Mézard conjecture for potentially Barsotti-Tate representations.* Forum of Mathematics Pi, 2 (2014), e1, 56 pp.

[38] I. M. Gelfand; I. Piatetski-Shapiro: *Automorphic forms and representation theory.* Trudy Moscow. Mat. Obšč. 12 (1963), 389-412.

[39] F. Q. Gouvêa: *Deformations of Galois representations.* IAS/Park City Mathematics Series, Vol. 9, 2001.

[40] A. Granville; T. J. Tucker: *It's as easy as abc.* Notices of the American Mathematical Society, 2009, 49 (10): 1224-1231.

[41] B. H. Gross: *On the Satake isomorphism.* Galois representations in arithmetic algebraic geometry (Durham, 1996), 223-237, London Math. Soc. Lecture Note Ser., 254, Cambridge Univ. Press, Cambridge, 1998.

[42] A. Grothendieck: *Technique de descente et théorèmes d'existence en géomérie algébrique. II. Le théorème d'existence en théorie formelle des modules* Séminaire Bourbaki, Vol. 5 (Paris), Soc. Math. France, 1995, Exp. No. 195, 369-390.

[43] Harish-Chandra: *Automorphic forms on a semisimple Lie group.* Proc. Nat. Acad. Sci. USA 45 (1959), 570-573.

[44] Harish-Chandra: *Automorphic forms on semisimple Lie groups.* Lecture Notes in Mathematics, vol. 62, Springer-Verlag, Berlin, 1968.

[45] H. Jacquet, R. Langlands: *Automorphic forms on* GL(2). Lecture Notes in Mathematics, vol. 114, Springer-Verlag, Berlin, 1970.

[46] H. Jacquet, J. A. Shalika: *On Euler products and the the classification of automorphic representations I.* Amer. J. Math. 103, 1981, 499-558.

[47] C. Khare: *Serre's modularity conjecture: a survey of the level one case.* in *L-functions and Galois representations*, London Math. Soc. Lecture Note Ser., vol. 320, Cambridge Univ. Press, Cambridge, 2007, 270-299.

[48] C. Khare; J-P. Wintenberger: *Serre's modularity conjecture. 1* Invent. Math. 178 (2009), no. 3, 485-504.

[49] C. Khare; J-P. Wintenberger: *Serre's modularity conjecture. 2* Invent. Math. 178 (2009), no. 3, 505-586.

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

[51] M. Kisin: *The Fontaine-Mazur conjecture for GL_2.* J.A.M.S. 22(3) (2009) 641-690.

[52] M. Kisin: *The structure of potentially semi-stable deformation rings.* Proceedings of ICM 2010, Vol. II, 294-311.

[53] A. W. Knapp: *Introduction to the Langlands program.* In Representation Theory and Automorphic Forms, Proceedings of Symposia in Pure Mathematics, vol. 61, 1997.

[54] A. W. Knapp; D. A. Vogan: *Cohomological induction and unitary representations.* Princeton University Press, Princeton, NJ, 1995.

[55] R. P. Langlands: *Base change for GL_2*. Annales of Math. Studies, vol. 96, Princeton University Press, Princeton, N. J., 1980.

[56] H. Maass: *Über eine neue Art von nicht analytischen automorphen Funktionene und die Bestimmung Reihen durch Functionalgleichungen.* Math. Annalen 121 (1949), 141-183.

[57] J. Manning: *Patching and multiplicity 2^k for Shimura curves.* Preprint. https://arxiv.org/abs/1902.06878

[58] J. Manning, J. Shotton: *Ihara's lemma for Shimura curves over totally real fields via patching.* Prepring. https://arxiv.org/abs/1907.06043

[59] B. Mazur: *Modular curves and the Eisenstein ideal.* Publ. Math. IHES 47, 33-186(1977).

[60] B. Mazur: *An introduction to the deformation theory of Galois representations.* in *Modular forms and Fermat's last theorem* G. Cornell, J. H. Silverman, G. Stevens, Springer.

[61] J. Nekovář: *On the parity of ranks of Selmer groups.* Compos. Math. 145 (2009), no. 6, 1351-1359.

[62] P. Ribenboim: *13 Lectures on Fermat's Last Theorem.* Springer-Verlag, New York and Heidelberg, 1979.

[63] K. Ribet: *On modular representations of $Gal(\overline{Q}/Q)$ arising from modular forms.* Invent. Math. 100 (1990), no. 2, 431-476.

[64] K. Rubin: *Modularity of mod 5 Galois representations.* in *Modular forms and Fermat's last theorem* G. Cornell, J. H. Silverman, G. Stevens, Springer.

[65] H. Saito: *Automorphic forms and extensions of number fields.* Lectures in Math., No. 8, Kinokuniya Book Store Co. Ltd., Tokyo, Japan, 1975.

[66] D. Savitt: *On a conjecture of Conrad, Diamond, and Taylor.* Duke Math. J. 128 (2005), no. 1, 141-197.

[67] M. Schlessinger: *Functors of Artin rings.* Trans. A. M. S. 130 (1968), 208-222.

[68] P. Scholze: *On the p-adic cohomology of the Lubin-Tate tower.* Ann. Sci. Éc. Norm. Supér. (4) 51, 2018, 811-863, with an appendix by Michael Rapoport.

[69] J.-P. Serre: *Sur les représentations modulaires de degré* 2 *de* $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Duke Mathematical Journal, 1987, 259-331.

[70] J.-P. Serre: *Propriétés galoisiennes des points d'ordre fini des courbes elliptiques.* Invent. Math. 15 (1972), 1987, 54 (1), 179-230.

[71] I. Shafarevich: *Algebra I.* Encyclopaedia of Mathematical Sciences, vol. 11, Springer-Verlag, 1990.

[72] J. A. Shalika: *The multiplicity one theorem for* GL_n. Annals of Math. 100, 1974, 171-193.

[73] G. Shimura: *Introduction to the arithmetic theory of automorphic functions.* Princeton University Press, Princeto, NJ, 1971.

[74] T. Shintani: *On liftings of holomorphic cusp forms.* in Proc. Symp. Pure Math., Vol. 33, Part 2, AMS, Providence, 1979, 97-110.

[75] E. de. Shalit: *Hecke rings and universal deformation rings.* in *Modular forms and Fermat's last theorem* G. Cornell, J. H. Silverman, G. Stevens, Springer.

[76] C. L. Siegel: *Topics in complex function theory.* vol. 3, Wiley-Interscience, New York, 1973.

[77] J. H. Silverman: *The arithmetic of elliptic curves.* Graduate Texts in Mathematics 106, Springer.

[78] J. H. Silverman: *Advanced topics in the arithmetic of elliptic curves.* Graduate Texts in Mathematics 151, Springer.

[79] S. Singh: *Fermat's Last Theorem.* ISBN 978-1857025217 (1997).

[80] G. Stevens: *An overview of the proof of Fermat's Last Theorem.* in *Modular forms and Fermat's last theorem* G. Cornell, J. H. Silverman, G. Stevens, Springer.

[81] R. Taylor: *On Galois representations associated to Hilbert modular forms.* Invent. Math. 98 (1989), 265-280.

[82] R. Taylor: *Remarks on a cojecture of Fontaine and Mazur.* J. Inst. Math. Jussieu 1 (2002), no. 1, 125-143.

[83] R. Taylor: *On icosahedral Artin representations II.* Amer. J. Math. 125 (2003), no. 3, 549-566.

[84] R. Taylor: *On the meromorphic continuation of degree two L-functions.* Doc. Math. 2006, Extra Vol., 729-779 (electronic).

[85] R. Taylor: *Automorphy for some l-adic lifts of automorphic mod l Galois representations. II.* Publ. Math. Inst. Hautes Études Sci. 108: 183-239.

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

[87] J. Tunnell: *Artin's conjecture for representations of octahedral type.* Bull. Amer. Math. Soc. 5 (1981), no. 2, 173-175.

[88] P. Vojta: *Diophantine approximations and value distribution theory.* Lect. Notes in Math. 1239, (1987).

[89] L. C. Washington: *Galois Cohomology.* in *Modular forms and Fermat's last theorem* G. Cornell, J. H. Silverman, G. Stevens, Springer.

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

[91] The Stack Project. `https://stacks.math.columbia.edu/`

[92] 模形式和费尔马大定理. First part of these notes.

[93] *Galois representations and their deformations.* Second part of these notes.