AUTOMORPHY LIFTING

LECTURES BY RICHARD TAYLOR,
NOTES BY DAN DORE AND TONY FENG

CONTENTS

1. THE LANGLANDS CORRESPONDENCE

1.1. **Introduction.** The notion of "automorphy" refers to automorphic forms – we say that a geometric object, a Galois representation, etc. is *automorphic* if its L-function agrees with the L-function of an automorphic form. When such results are available, they are extremely powerful. For example, if an L-function agrees with that of an automorphic form, we can construct a meromorphic continuation with the expected sort of functional equation to the entire complex plane. These methods are behind the proof of Fermat's Last Theorem, the Sato-Tate conjecture, and many more things.

However, it was not until the work of Wiles in the 1990's that we had many results at all establishing automorphy of interesting Galois representations. The principal tool here is *automorphy lifting*. The basic principle is that we consider a Galois representation $\rho\colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Q}_\ell)$ and consider the reduction mod ℓ, i.e. $\overline{\rho}\colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{F}_\ell)$. In many cases of interest, we can define and prove a notion of automorphy for these "mod ℓ" representations. Then, we prove a "lifting" theorem that allows us to show that this implies automorphy for the "mod ℓ^n" representations for each n. Since the $\mathbf{Z}_\ell$-representation is *defined* as the inverse limit of the $\mathbf{Z}/\ell^n$ representations (and the $\mathbf{Q}_\ell$ representation is obtained from this by inverting ℓ), this gives us automorphy of the full thing.

1.2. **The global Langlands correspondence for** GL_n**.** In the next sections we will explain the following conjecture, which is basic to the whole subject.

Conjecture 1.2.1 (Langlands Correspondence for GL_n over F)**.** *Let F be a number field. Fix an identification $\iota\colon \overline{\mathbf{Q}_\ell} \simeq \mathbf{C}$. Then there exists a bijection between the following two sets:*

- *Algebraic cuspidal automorphic representations $\pi = \otimes'_v \pi_v$ of $\mathrm{GL}_n(\mathbf{A}_F)$.*
- *Irreducible algebraic ℓ-adic representations r of $G_F := \mathrm{Gal}(\overline{F}/F)$.*

Furthermore, this bijection satisfies the following compatibilities with the local Langlands correspondences at each finite place.

(1) *For each finite place v of F, the local Langlands correspondence relates π_v to $r|_{G_{F_v}}$. More precisely, we require:*

$$\mathrm{rec}(\pi_v) = \mathrm{WD}(r|_{G_{F_v}}).$$

(2) *If $\tau\colon F \hookrightarrow \overline{\mathbf{Q}}_\ell$ is given, then $\mathrm{HT}_\tau(r) = -\mathrm{HC}_{\iota\circ\tau}(\pi_\infty)$. Here,* HT *denotes the* Hodge-Tate weights *and* HC *denotes the* Harish-Chandra weights.

Remark 1.2.2. The two sets have the same cardinality for trivial reasons, so the compatibility with the local theory is *crucial.* This is what allows us to show an equality of L-functions!

1.3. **Automorphic forms.** Now, we will define the various terms appearing in Conjecture 1.2.1. The *ring of adeles* $\mathbf{A}_F$ is defined as the restricted product $\prod'_v F_v$, and we topologize $\mathrm{GL}_n(\mathbf{A}_F)$ by viewing it as an open subset of $\mathbf{A}_F^{n^2+1}$ (the last coordinate is $\det^{-1}$ – this is like how we topologize the ideles by the embedding $x \mapsto (x, x^{-1})$ of $\mathbf{A}^\times$ into $\mathbf{A}^2$). We let $\mathbf{A}_F^\infty = \prod'_{v \text{ finite}} F_v$ and $F_\infty = \prod_{v|\infty} F_v$, so $\mathbf{A}_F = \mathbf{A}_F^\infty \times F_\infty$ and

$$\mathrm{GL}_n(\mathbf{A}_F) = \mathrm{GL}_n(F_\infty) \times \mathrm{GL}_n(\mathbf{A}_F^\infty).$$

Now, we can define the space of *cuspidal automorphic forms* for GL_n:

Definition 1.3.1. The space $\mathscr{A}_0(\mathrm{GL}_n(F)\backslash \mathrm{GL}_n(\mathbf{A}_F))$ of *cuspidal automorphic forms* for GL_n is defined to be the set of functions $\varphi\colon \mathrm{GL}_n(F)\backslash\mathrm{GL}_n(\mathbf{A}_F) \to \mathbf{C}$ satisfying the following conditions:

(i) φ is *smooth*, i.e. there is an open subset $W \subseteq \mathrm{GL}_n(\mathbf{A}_F^\infty)$ such that the restriction of φ to $W \times \mathrm{GL}_n(F_\infty)$ factors through the projection to $\mathrm{GL}_n(F_\infty)$ via a *smooth* map between the manifolds $\mathrm{GL}_n(F_\infty)$ and $\mathbf{C}$. In other words, φ is locally constant in the finite places and smooth in the infinite places.

(ii) φ is $\mathrm{GL}_n(\widehat{\mathscr{O}_F}) \times U_\infty$-finite. In other words, the space of right translates of φ under $\mathrm{GL}_n(\widehat{\mathscr{O}_F}) \times U_\infty$ is *finite-dimensional.* Here, $\widehat{\mathscr{O}_F}$ is the profinite completion of $\mathscr{O}_F$, i.e.

$$\widehat{\mathscr{O}_F} = \prod_{v \text{ finite}}{}' \mathscr{O}_{F,v},$$

and

$$U_\infty = \prod_{v|\infty} U_v,$$

where U_v is a maximal compact subgroup of $\mathrm{GL}_n(F_v)$, required to be $\mathrm{O}(n)$ when v is real and $\mathrm{U}(n)$ when v is complex.

(iii) φ is $\mathfrak{z}$-finite (in the same sense of finite as in the previous condition). Here, $\mathfrak{z}$ is the center of the universal enveloping algebra $\mathscr{U}$ of the Lie algebra $\mathfrak{g} := (\mathrm{Lie}\,\mathrm{GL}_n(F_\infty)) \otimes \mathbf{C}$, and $\mathrm{Lie}\,\mathrm{GL}_n(F_\infty)$ acts on φ by the formula:

$$(X\varphi)(g) = \frac{d}{dt}\varphi(g\exp(tX))\,|_{t=0}$$

We extend the action to $\mathfrak{g}$ by $\mathbf{C}$-linearity and then to $\mathscr{U}$ (by the universal property of $\mathscr{U}$ - i.e. it's the universal associative algebra containing $\mathfrak{g}$).

Note that

$$\mathfrak{g} = \mathrm{Mat}_{n\times n}(F_\infty) \otimes_{\mathbf{R}} \mathbf{C} = \mathrm{Mat}_{n\times n}(F \otimes_{\mathbf{Q}} \mathbf{C}) = \prod_{\tau\colon F \hookrightarrow \mathbf{C}^1} M_{n\times n}(\mathbf{C}) =: \prod_{\tau\colon F\hookrightarrow \mathbf{C}} \mathfrak{g}_\tau.$$

We similarly have $\mathscr{U} = \otimes_\tau \mathscr{U}_\tau$ and $\mathfrak{z} = \otimes_\tau \mathfrak{z}_\tau$.

Harish-Chandra showed that $\mathfrak{z}_\tau \simeq \mathbf{C}[x_1,\dots,x_n]^{S_n}$. This isomorphism is characterized by the fact that on the irreducible representation of GL_n with highest weight given by $a_1 \geq a_2 \geq \cdots \geq a_n$, then the corresponding character of $\mathfrak{z}_\tau$ sends

$$\{x_1,\dots,x_n\} \mapsto \{a_1 + \frac{n-1}{2}, a_2 + \frac{n-3}{2}, \cdots, a_n + \frac{1-n}{2}\}.$$

(iv) φ is *slowly increasing*. This means that there exist real constants C, s such that

$$|\varphi(g)| \leq C\|g\|^s \quad \text{for all} \quad g \in \mathrm{GL}_n(\mathbf{A}_F).$$

Here, for $g = (g_v \in \mathrm{GL}_n(F_v))_v$, we define the norm

$$\|g\| := \prod_v \max_{1\leq i,j\leq n}\{|(g_v)_{i,j}|_v, |(g_v)^{-1}_{i,j}|_v\}$$

where $(g_v)_{i,j}$ are the matrix coordinates.

(v) φ is *cuspidal*. We define the block upper-triangular unipotent subgroup

$$N_m = \left\{\begin{pmatrix} 1_m & * \\ 0 & 1_{n-m}\end{pmatrix}\right\}$$

for $0 < m < n$. As m varies, we run through the set of unipotent radicals of representatives of each conjugacy class of maximal parabolic subgroups.

The condition that φ is cuspidal says that for each $0 < m < n$ and each $g \in \mathrm{GL}_n(\mathbf{A}_F)$, we have:

$$\int_{N_m(F)\backslash N_m(\mathbf{A}_F)} \varphi(ug)\,du = 0$$

Remark 1.3.2. Sometimes one wants to discuss automorphic forms which are not necessarily cuspidal (e.g. modular forms which are not cusp forms). These are essentially functions on $\mathrm{GL}_n(F)\backslash\mathrm{GL}_n(\mathbf{A}_F)$ satisfying conditions (i) to (iv) above, except that one often wants to strengthen the growth condition (iv) in this case.

1.4. **Automorphic representations.** The space $\mathscr{A}_0(\mathrm{GL}_n(F)\backslash \mathrm{GL}_n(\mathbf{A}_F))$ is not *quite* a representation of the group $\mathrm{GL}_n(\mathbf{A}_F)$. This is due to the fact that if φ is U_∞-finite, then the translate of φ by $g \in \mathrm{GL}_n(F_\infty)$ is $gU_\infty g^{-1}$-finite, and not necessarily U_∞-finite. However, it admits an action of $\mathrm{GL}_n(\mathbf{A}_F^\infty) \times U_\infty$ as well as an action of $\mathfrak{g}$. These are related by the fact that the differential of the U_∞ action is the restriction of the $\mathfrak{g}$-action to $\mathrm{Lie}(U_\infty)$, i.e.

$$g(X\varphi) = (\mathrm{ad}(g_\infty)(X))(g\varphi) \text{ for } g \in \mathrm{GL}_n(\mathbf{A}_F^\infty) \times U_\infty \text{ and } X \in \mathfrak{g}.$$

By replacing the space $\mathscr{A}_0(\mathrm{GL}_n(F)/\mathrm{GL}_n(\mathbf{A}_F))$ by a Hilbert space completion, we can get an honest representation of $\mathrm{GL}_n(\mathbf{A}_F)$. However, we then have to worry about the topology of the vector space, so it is a trade-off.

1.4.1. Now look back at Conjecture 1.2.1. On one side of this correspondence, we have the set of algebraic cuspidal automorphic representations $\pi = \bigotimes_v' \pi_v$ of $\mathrm{GL}_n(\mathbf{A}_F)$. These live inside the space $\mathscr{A}_0(\mathrm{GL}_n(F)\backslash \mathrm{GL}_n(\mathbf{A}_F))$ of cuspidal automorphic forms, which carries a canonical structure of a $(\mathrm{GL}_n(\mathbf{A}_F^\infty) \times U_\infty, \mathfrak{g})$-module. This space decomposes discretely as $\bigoplus \pi$ with π ranging through the set of irreducible cuspidal automorphic representations, each appearing exactly once (this is the "Multiplicity One" theorem, which is true for GL_n but not for other reductive groups).

1.4.2. Let π be one of these irreducible pieces. We can write $\pi = \bigotimes_v' \pi_v$. For each finite v, π_v is an irreducible representation of $\mathrm{GL}_n(F_v)$, which is also *smooth* and *admissible*. The condition of being smooth means that the stabilizer of every vector is open, and the condition of being admissible means that the space of vectors fixed by a given non-empty open subset of $\mathrm{GL}_n(F_v)$ is finite-dimensional. For each infinite v, π_v is an irreducible *admissible* $(\mathfrak{g}_v, U_v)$-module. Here, admissibility means that the U_v-isotypic components of π_v are finite dimensional.

What does it mean to say that $\pi = \bigotimes_v' \pi_v$? How do we make sense of an infinite tensor product? For all but finitely many v[2], the representation π_v is *unramified* (or "spherical") meaning that the space of $\mathrm{GL}_n(\mathscr{O}_v)$-fixed points $\pi_v^{\mathrm{GL}_n(\mathscr{O}_v)}$ is non-zero. In fact, this implies (by commutativity of the local spherical Hecke algebra $\mathrm{GL}_n(\mathscr{O}_v)\backslash \mathrm{GL}_n(F_v)/\mathrm{GL}_n(\mathscr{O}_v)$) that $\pi_v^{\mathrm{GL}_n(\mathscr{O}_v)}$ is one-dimensional. We choose (non-canonically!) a basis vector e_v for this space for each v. Now, we can define the infinite tensor product:

$$\bigotimes_v{}' \pi_v := \varinjlim_{\substack{S \\ \#S<\infty \\ S \supset \text{all archimedean/ramified } v}} \bigotimes_{v \in S} \pi_v.$$

The maps in the direct system are defined for each $S \subseteq T$ by the map:

$$\bigotimes_{v\in S} \pi_v \to \bigotimes_{v \in T} \pi_v : \otimes_{v\in S} x_v \mapsto (\otimes_{v \in S} x_v) \otimes (\otimes_{v \in T-S} e_v).$$

Varying the choices of e_v amounts to scaling these maps, so we get an isomorphic infinite tensor product.

[2]We will abbreviate this as "almost all v" throughout.

1.4.3. Now, the center of the universal enveloping algebra of $\mathrm{GL}_n(F_\infty)$ is $\mathfrak{z} = \bigotimes_{\tau : F \hookrightarrow \mathbf{C}} \mathfrak{z}_\tau$, and this acts on π by scalars. There's an isomorphism $\mathfrak{z}_\tau \simeq \mathbf{C}[x_1, \dots, x_n]^{S_n}$, so these scalars give us a multiset of n complex numbers, called the *Harish-Chandra parameters* of π_∞ and denoted $\mathrm{HC}_\tau(\pi_\infty)$. This allows us to make the following definition.

Definition 1.4.1. We say that π is *algebraic* if $\mathrm{HC}_\tau(\pi_\infty) \subseteq \mathbf{Z}$ for each τ. We say that π is *regular* if $\mathrm{HC}_\tau(\pi_\infty)$ has n *distinct* elements for each τ.

Remark 1.4.2. The condition that π is regular and algebraic is equivalent to the condition that π is *cohomological*, which means that it appears in the Betti cohomology of the symmetric space $\mathrm{GL}_n(F) \backslash \mathrm{GL}_n(\mathbf{A}_F)/K$ for some choice of compact subgroup K (e.g. for $n = 2$ and $F = \mathbf{Q}$, this is a modular curve, with level specified by the choice of K). This means that the regular algebraic representations are accessible via *topology*. In general, very little can be said if we do not make these assumptions on π.

1.5. **The Galois side.** Now, we have defined the objects appearing on the "automorphic" side of the global Langlands correspondence. On the other side, we have the "Galois" side. We consider the profinite absolute Galois group

$$G_F := \mathrm{Gal}(\overline{F}/F) = \varprojlim_{[E:F]<\infty} \mathrm{Gal}(E/F)$$

and its ℓ-adic representations (for ℓ some prime number). These are defined to be *continuous* homomorphisms $r : G_F \to \mathrm{GL}_n(\overline{\mathbf{Q}}_\ell)$.

The group G_F as an abstract profinite group is extremely complicated - for example, by work of Thompson, the monster group is a quotient in infinitely many ways! However, we can consider the *decomposition subgroups* $G_{F_v} = \mathrm{Gal}(\overline{F_v}/F_v) \hookrightarrow G_F$. These are well-defined only up to conjugation by G_F (the injection is determined by choosing a place of $\overline{F}$ over v, and G_F permutes these). These decomposition groups are much simpler - they are *pro-solvable*. Thus, the representation theory of G_F is more manageable if we take into account this family of subgroups.

1.5.1. *The Weil group.* It turns out to be convenient to replace the decomposition group F_v by the *Weil group* W_{F_v}. To define this, we consider the *geometric Frobenius* automorphism Frob_v of the residue field $k(v)$ at v. This is the *inverse* of the map $\alpha \mapsto \alpha^{\#k(v)}$ (the latter is called the *arithmetic Frobenius*). There is an isomorphism

$$G_{k(v)} \xrightarrow{\sim} \widehat{\mathbf{Z}} = \varinjlim_m \mathbf{Z}/m\mathbf{Z}$$

which sends Frob_v to 1 (we say that $G_{k(v)}$ is *pro-cyclic* with generator Frob_v). An automorphism of $\overline{F}_v$ induces an automorphism of the residue field $\overline{k(v)}$, so we have a canonical surjection $G_{F_v} \to G_{k(v)}$. The kernel is called the *inertia group* I_v. Thus, we have a short exact sequence (which splits, but non-canonically):

$$0 \longrightarrow I_v \longrightarrow G_{F_v} \longrightarrow G_{k(v)} \longrightarrow 0$$

By modifying this exact sequence slightly, we obtain the *Weil group*:

$$0 \longrightarrow I_v \longrightarrow W_{F_v} \xrightarrow{v} \mathrm{Frob}_v^{\mathbf{Z}} \simeq \mathbf{Z} \longrightarrow 0$$

In other words, the Weil group is the subgroup of G_{F_v} which acts by integer powers of the Frobenius on $\overline{k(v)}$. We topologize it via the above exact sequence: we require that I_v is an *open* subgroup, so W_{F_v} is homeomorphic to $I_v \times \mathbf{Z}$ with the discrete topology on $\mathbf{Z}$. The map $v : W_{F_v} \to \mathbf{Z}$ is called the *valuation map*.

1.5.2. *Weil-Deligne representations.*

Definition 1.5.1. Let V be a vector space over a field L of characteristic 0.[3] A *Weil-Deligne representation* (abbreviated "WD-representation") of W_{F_v} on V is a pair (ρ, N) with $\rho : W_{F_v} \to \mathrm{GL}(V)$ a map with *open kernel* (i.e. it is continuous with respect to the *discrete topology* on V) and N an endomorphism of V which satisfies:

$$\rho(\sigma)N\rho(\sigma)^{-1} = \#k(v)^{-v(\sigma)}N$$

for all $\sigma \in W_{F_v}$. This condition implies that N is *nilpotent* (by examining its eigenvalues).

This definition is somewhat strange, but it allows us to ignore the topology of V. This means that the notion is the same over $\mathbf{C}$ and over $\overline{\mathbf{Q}}_\ell$. In addition, there are only finitely many choices of N (up to isomorphism) for a fixed ρ, so the N can be thought of as an additional finite bit of combinatorial data.

Definition 1.5.2. We say that a WD-representation (ρ, N) is *Frobenius-semisimple* if $\rho(\Phi)$ is semisimple (i.e. it is diagonalizable over $\overline{L}$) for any lift $\Phi \in W_{F_v}$ of Frob_v. This is equivalent to requiring that ρ be semisimple in the abelian category of WD representations.

We say that (ρ, N) is *semisimple* if it is Frobenius-semisimple and $N = 0$.

Proposition 1.5.3. *Given a WD-representation (ρ, N), there exists a unique unipotent $u \in \mathrm{GL}(V)$ which commutes with N and $\mathfrak{I}(\rho)$ and such that $(\rho u^{-v(\sigma)}, N)$ is Frobenius-semisimple. We call this the* Frobenius semi-simplification *of (ρ, N).*

Proposition 1.5.4 (Local Langlands Correspondence)**.** *For a finite place v, there exists a natural bijection:*

$$\left\{\begin{matrix}\textit{irred. smooth rep'ns} \\ \textit{of}\,\mathrm{GL}_n(F_v)\textit{ over}\,\mathbf{C}\end{matrix}\right\} \xrightarrow{\mathrm{rec}} \left\{\begin{matrix}n\textit{-dim Frob-semisimplel} \\ \textit{WD-rep'ns} \\ \textit{of}\, W_{F_v}\textit{ over}\,\mathbf{C}\end{matrix}\right\}.$$

Note that we need to say what "natural" means for this to have much meaning (in particular, the L-functions and ϵ-factors on both sides match, there is compatibility between the various n, etc.).

Example 1.5.5. For $n = 1$, this is essentially the content of local class field theory: the Artin map gives an isomorphism

$$\mathrm{Art}_{F_v} : F_v^\times \xrightarrow{\sim} W_{F_v}^{\mathrm{ab}}$$

which sends uniformizers to lifts of Frob_v. Then we have $\mathrm{rec}(\chi) = \chi \circ \mathrm{Art}_{F_v}^{-1}$ for χ a character (i.e. a 1-dimensional representation) of $F_v^\times$.

[3]the characteristic 0 condition is not technically relevant to the definition, but all known applications are in this setting.

1.6. **ℓ-adic representations.** We have almost finished explaining Conjecture 1.2.1. It remains to discuss "irreducible algebraic ℓ-adic representations".

Definition 1.6.1. An *ℓ-adic representation* $G_F \to \mathrm{GL}_n(\overline{\mathbf{Q}}_\ell)$ is a continuous representation (with the profinite topology on the domain, and the ℓ-adic topology on the target). We say that an ℓ-adic representation $r: G_F \to \mathrm{GL}_n(\overline{\mathbf{Q}}_\ell)$ is *algebraic* (which is often called "geometric" in the literature) if:

(1) For almost all v, r is *unramified* at v, meaning $r(I_{F_v}) = \{1\}$.
(2) For all $v \mid \ell$, $r|_{G_{F_v}}$ is "de Rham" (a notion that we will take as a black box). We emphasize that this is a *purely local* constraint.

Remark 1.6.2. Any representation that comes from the ℓ-adic cohomology of algebraic varieties satisfies these two conditions. Fontaine-Mazur conjectured that the converse is true.

1.6.1. *The associated Weil-Deligne representation.* Suppose $v \nmid \ell$. Fix $\phi \in W_{F_v}$ lifting Frobenius. Then there is a map

$$t: I_{F_v} \twoheadrightarrow \mathbf{Z}_\ell,$$

which is unique up to multiplication by $\mathbf{Z}_\ell^\times$. Then there exists a unique WD representation (ρ, N) such that

$$r|_{W_{F_v}}(\sigma) = \rho(\sigma)\exp(t_\ell(\phi^{-v(\sigma)}\sigma)N) \quad \text{for all } \sigma \in W_{F_v}. \tag{1.6.1}$$

What's surprising about this? The representation r is *continuous* with respect to the profinite topology on G_{F_v} and the ℓ-adic topology on $\mathrm{GL}_n(\overline{\mathbf{Q}}_\ell)$, while ρ is continuous with respect to the profinite topology on I_{F_v} and the *discrete* topology on $\mathrm{GL}_n(\overline{\mathbf{Q}}_\ell)$. So the conversion (1.6.1) strips out the ℓ-adic topology.

This construction appears to depend on the choices of ϕ and t, but up to isomorphism it is actually independent of these choices. Since the notions of $\mathbf{C}$ and $\overline{\mathbf{Q}}_\ell$-valued WD representations are identified by the choice of ι, this provides a translation between WD representations over $\mathbf{C}$ and $\overline{\mathbf{Q}}_\ell$.

Remark 1.6.3. For $v \mid \ell$, there is still a way to define a good notion of Weil-Deligne representation $\mathrm{WD}(r|_{G_{F_v}})$, but it no longer determines $r|_{G_{F_v}}$.

1.6.2. *Hodge-Tate weights.*

Definition 1.6.4. The *cyclotomic character*

$$\epsilon_\ell: G_F \to \mathbf{Z}_\ell^\times$$

is characterized by the property that for $\sigma \in G_F$ and ζ a ℓ-power root of unity,

$$\sigma(\zeta) = \zeta^{\epsilon_\ell(\sigma)}.$$

This is continuous and algebraic, and unramified away from ℓ. Furthermore, for $v \mid \ell$ we have

$$\epsilon_\ell(\mathrm{Frob}_v) = \# k(v)^{-1}.$$

For $\tau\colon F \hookrightarrow \overline{\mathbf{Q}}_\ell$, the closure of the image is F_v for some $v \mid \ell$. We want to define its set of Hodge-Tate weights, $\mathrm{HT}_\tau(r)$. To do this, we take $r|_{G_{F_v}}$ and tensor it with a power of the cyclotomic character ϵ_ℓ^i, and extend scalars:

$$(r|_{G_{F_v}}\epsilon_\ell^i \otimes_{\tau,F_v} \overline{F}_v)$$

which has an action of $\overline{\mathbf{Q}}_\ell \otimes_{\tau,F_v} \overline{F}_v$. This has a semi-linear action of G_{F_v}, via r on the first factor and the Galois action on the second. Hence we can take G_{F_v}-fixed points. Tate proved that

$$\dim_{\overline{\mathbf{Q}}_\ell}(r|_{G_{F_v}}\epsilon_\ell^i \otimes_{\tau,F_v} \overline{F}_v)^{G_{F_v}} \le n \tag{1.6.2}$$

(Equality holds for algebraic r.) Then $\mathrm{HT}_\tau(r)$ is a multiset of integers in which $i \in \mathbf{Z}$ has multiplicity equal to $\dim_{\overline{\mathbf{Q}}_\ell}(r|_{G_{F_v}}\epsilon_\ell^i \otimes_{\tau,F_v} \overline{F}_v)^{G_{F_v}}$.

Thus for $v \mid \ell$, we have attached "discrete" invariants $\mathrm{WD}(r|_{G_{F_v}})$ and $\mathrm{HT}_\tau(r)$. These don't determine everything – there are continuous families of possible r.

Both the analytic and Galois-theoretic world are very big, but the algebraic stuff inside them looks kind of discrete (and this discrete stuff should be related to motives defined over F).

1.7. **What is known?**

1.7.1. *The case $n = 1$.* There is one case where Conjecture 1.2.1 is entirely known, namely the $n = 1$ case. Let me explain the association $\chi \mapsto r_{\ell,\iota}(\chi)$. In this case there is no complication at ∞, so a cuspidal automorphic representation of GL_1 (what we denoted by π above) is the same as a continuous character

$$\chi\colon \mathbf{A}_F/F^\times \to \mathbf{C}^\times.$$

The "algebraic condition" says that $\chi|_{(F_\infty^\times)^0}$, note that $(F_\infty^\times)^0$ is a product of copies of $\mathbf{C}^\times$ and $\mathbf{R}^\times$, looks like

$$x \mapsto \prod_{\tau\colon F\hookrightarrow\mathbf{C}} \tau(x)^{-n_\tau}.$$

The Harish-Chandra parameters are then $\mathrm{HC}_\tau(\chi) = \{-n_\tau\}$.

Remark 1.7.1. The continuous characters $\chi\colon \mathbf{A}_F/F^\times \to \mathbf{C}^\times$ are called *Grossencharacters.*

We were going to explain how to produce the 1-dimensional algebraic ℓ-adic representation r_χ. First we'll define another $\widetilde{\chi}\colon \mathbf{A}_F^\times \to \mathbf{C}^\times$, given by

$$\widetilde{\chi} := \chi(x) \prod_{\tau\colon F\hookrightarrow\mathbf{C}} \tau(x_\infty)^{n_\tau}. \tag{1.7.1}$$

What properties does $\widetilde{\chi}$ have? It is no longer trivial on $F^\times$. However, it does at least take $F^\times$ into $\overline{\mathbf{Q}}^\times$. The continuity of χ implies that it is invariant by some open compact subgroup $U \subset \mathbf{A}_F^\times$, so $\widetilde{\chi}$ is invariant by $U \cdot (F_\infty^\times)^0$. Now, a fundamental fact is that any such quotient

$$\mathbf{A}_F^\times/F^\times U \cdot (F_\infty^\times)^0$$

is actually finite. So the $\widetilde{\chi}$ is actually valued in $\overline{\mathbf{Q}}^\times$, even on $\mathbf{A}_F^\times$.

We then use $\iota^{-1}\circ\widetilde{\chi}\colon \overline{\mathbf{Q}}^\times \hookrightarrow \overline{\mathbf{Q}}_\ell^\times$. We want it to be invariant under $F^\times$, and we have the freedom to modify the component at ℓ. Therefore, we write down

$$\chi^{(\ell)}\colon \mathbf{A}_F^\times/F^\times \to \overline{\mathbf{Q}}_\ell^\times$$

by

$$x \mapsto \iota^{-1}\circ\widetilde{\chi}(x)\prod_{\tau\colon F\hookrightarrow\overline{\mathbf{Q}}_\ell}(\tau x_\ell)^{-n_{\iota\circ\tau}}. \tag{1.7.2}$$

Since this only involved the components above ℓ, it's still going to be invariant under $(F_\infty^\times)^0$, so the character will factor through

$$\chi^{(\iota)}\colon \mathbf{A}_F^\times/\overline{F^\times(F_\infty^\times)^0} \to \overline{\mathbf{Q}}_\ell^\times.$$

Theorem 1.7.2. *We have commutative diagrams*

$$\begin{tikzcd} \mathrm{Art}_F\colon \mathbf{A}_F^\times/\overline{F^\times(F_\infty^\times)^0} \arrow[r, "\sim"] & G_F^{\mathrm{ab}} \\ \mathrm{Art}_{F_v}\colon F_v^\times \arrow[u] \arrow[r, "\sim"] & W_{F_v}^{\mathrm{ab}} \arrow[u] \end{tikzcd}$$

for $v\nmid\infty$ and

$$\begin{tikzcd} \mathrm{Art}_F\colon \mathbf{A}_F^\times/\overline{F^\times(F_\infty^\times)^0} \arrow[r, "\sim"] & G_F^{\mathrm{ab}} \\ \mathrm{Art}_{F_v}\colon F_v^\times/(F_v^\times)^0 \arrow[u] \arrow[r, "\sim"] & G_{F_v}^{\mathrm{ab}} \arrow[u] \end{tikzcd}$$

for $v\mid\infty$.

Here Art_F and Art_{F_v} are the global and local Artin maps. Finally, we set

$$r_{\ell,\iota}(\chi) = \chi^{(\iota)}\circ\mathrm{Art}_F^{-1}.$$

In summary, the content of the Langlands correspondence for $n=1$ is essentially the content of class field theory.

1.7.2. *Results for CM fields.* The conjecture has two directions: going from (a) automorphic representations to (b) Galois representations, or the other way. The course is going to be about going from (b) to (a), but I will first make some comments about going from (a) to (b); we do not currently have any way to get (b) $\to$ (a) without first knowing (a) $\to$ (b).

Beyond $n=1$, the only cases where we know anything are when F is a CM field.

Definition 1.7.3. A number field F is a *CM field* if there exists $c\in\mathrm{Aut}(F)$ such that for all $\tau\colon F\hookrightarrow\mathbf{C}$, we have $c\circ\tau=\tau\circ c$. (In other words, there's a well-defined complex multiplication no matter how you put F into $\mathbf{C}$.)[4]

Remark 1.7.4. Evidently the automorphism c must have order 2. We can define $F^+ = F^{\{1,c\}}$; then F^+ is totally real and $[F:F^+]=1$ or 2.

[4]Under this definition, totally real fields are also CM fields, although some people like to exclude this case.

Suppose that π is a *regular* algebraic cuspidal automorphic representation of $\mathrm{GL}_n(\mathbf{A}_F)$ (to recall the definition, see Definition 1.4.1).[5] It is then a theorem that there exists an ℓ-adic representation $r_{\ell,\iota}(\pi)\colon G_F \to \mathrm{GL}_n(\overline{\mathbf{Q}}_\ell)$ such that for $v \nmid \ell$,

$$\mathrm{rec}(\pi)^{\mathrm{ss}} \cong \iota\,\mathrm{WD}(r_{\ell,\iota}(\pi)|_{G_{F_v}})^{\mathrm{ss}}$$

and if π_v is unramified, then so is $\iota\,\mathrm{WD}(r_{\ell,\iota}(\pi)|_{G_{F_v}})$.[6] In particular, we know that $r_{\ell,\iota}(\pi)$ is unramified almost everywhere. However, we don't know that $r_{\ell,\iota}(\pi)|_{G_{F_v}}$ is de Rham.

1.7.3. *Self-dual representations.* We can do much better with more hypotheses. Suppose F is CM and π is regular algebraic. Suppose $\chi : \mathbf{A}_{F^+}^\times/(F^+)^\times \to \mathbf{C}^\times$ is a grossencharacter such that if F is complex then $\chi_v(-1)$ is independent of $v \mid \infty$, and

$$\pi^c \cong \pi^\vee \otimes (\chi \circ N_{F/F^+} \circ \det). \tag{1.7.3}$$

This is some sort of self-duality condition.

Then there exists $r_{\ell,\iota}(\pi)\colon G_F \to \mathrm{GL}_n(\overline{\mathbf{Q}}_\ell)$ an algebraic ℓ-adic representation such that

(1) For all v, $\mathrm{rec}(\pi_v) = \iota\,\mathrm{WD}(G_{F_v})^{F-\mathrm{ss}}$ (conjecturally this semisimplicity is automatic),
(2) For all $\tau : F \hookrightarrow \overline{\mathbf{Q}}_\ell$, $\mathrm{HC}_{\iota\circ\tau}(\pi_\infty) = -\mathrm{HT}_\tau(r_{\ell,\iota}(\pi))$.

1.7.4. *A remark on the proofs.* The idea in §1.7.3 is to use the Arthur-Selberg trace formula to move to a unitary group, which has a Shimura variety.

This doesn't work for §1.7.2. The Galois representations there are not found in the cohomology of any particular motive, but pieced together from representations mod ℓ^n coming from motives.

1.7.5. *Base change.* What is base change? Given a Galois representation of G_F, you can restrict to G_E for a field extension E/F, and then go back to an automorphic representation for $\mathrm{GL}_n(E)$. Is there a way to realize this directly at the level of automorphic representations?

This is very difficult, but sometimes possible. It was done by Langlands for GL_2 and generalized to GL_n by Arthur-Clozel.

Remark 1.7.5. Although the plausibility argument at the beginning was only for algebraic representations, the method works for all automorphic representations.

Theorem 1.7.6. *Suppose E/F is a finite solvable Galois extension of CM fields.*

(1) If π is a cuspidal automorphic representation of $\mathrm{GL}_n(\mathbf{A}_F)$, then there exists $n = n_1 + \ldots + n_r$, with each $n_i \in \mathbf{Z}_{>0}$, and cuspidal automorphic representations π_i of $\mathrm{GL}_{n_i}(\mathbf{A}_E)$, such that:

$$\mathrm{rec}(\pi_v)|_{W_{E_w}} \cong \bigoplus_i \mathrm{rec}(\pi_{i,w}) \textit{ for all } w \mid v,$$

and

$$\mathrm{HC}_\tau(\pi_\infty) = \coprod_i \mathrm{HC}_\tau(\pi_{\iota,\infty}) \textit{ for all } \tau : E \hookrightarrow \mathbf{C},$$

[5] We know almost nothing outside of the regular case.

[6] An unramified Weil-Deligne representation (ρ, N) is one with $N = 0$ and ρ unramified.

(2) *If r is a regular algebraic ℓ-adic representation $G_F \to \mathrm{GL}_n(\overline{\mathbf{Q}}_\ell)$ with $r|_{G_E}$ irreducible, then r is automorphic if and only if $r|_{G_E}$ is automorphic.*

(3) *If π is a cuspidal automorphic representation of $\mathrm{GL}_n(\mathbf{A}_E)$ then there exists $n_1, \ldots, n_r \in \mathbf{Z}_{>0}$ with $n_1 + \ldots + n_r = n[E:F]$ and cuspidal automorphic representations π_i of $\mathrm{GL}_{n_i}(\mathbf{A}_F)$*

$$\oplus_i \operatorname{rec}(\pi_{i,v}) = \bigoplus_{w|v} \operatorname{Ind}_{W_{E_w}}^{W_{F_v}} \operatorname{rec}(\pi_w)$$

and

$$\coprod_i \mathrm{HC}_\tau(\pi_{i,\infty}) = \coprod_{\substack{\tau\colon E\hookrightarrow\mathbf{C} \\ \tilde{\tau}|_F=\tau}} \mathrm{HC}_{\tilde{\tau}}(\pi_\infty) \text{ for all } \tau\colon E \hookrightarrow \mathbf{C},$$

(4) *If r is an algebraic irreducible $G_E \to \mathrm{GL}_n(\overline{\mathbf{Q}}_\ell)$ and $\operatorname{Ind}_{G_E}^{G_F} r$ is irreducible, then r is automorphic if and only if $\operatorname{Ind}_{G_E}^{G_F}$ is automorphic.*

The upshot for us is that we can freely make solvable base change to simplify our lives, which we will repeatedly do.

1.8. **Mod-ℓ representations.** We have said what it means for an ℓ-adic representation $r\colon G_F \to \mathrm{GL}_n(\overline{Q}_\ell)$ to be automorphic, but we want to introduce the notion of automorphy for a *mod ℓ representation*

$$\overline{r}\colon G_F \to \mathrm{GL}_n(\overline{\mathbf{F}}_\ell).$$

Note that any such $\overline{r}$ will automatically be unramified at almost all places. Also there is no longer a notion of “de Rham over $v \mid \ell$”, so there’s no obvious obstruction for all such $\overline{r}$ to be “algebraic” (what is commonly called “geometric”).

Any ℓ-adic representation $r\colon G_F \to \mathrm{GL}_n(\overline{\mathbf{Q}}_\ell)$ can be conjugated by some $g \in \mathrm{GL}_n(\overline{\mathbf{Q}}_\ell)$ such that

$$grg^{-1}\colon G_F \to \mathrm{GL}_n(\mathscr{O}_{\overline{\mathbf{Q}}_\ell}).$$

We then define the *reduction* of r to be

$$\overline{r} := (grg^{-1} \mod \mathfrak{m})^{\mathrm{ss}}.$$

Here the ss standards for semisimplification, and is necessary to make this well-defined (a priori it depends on g).

Definition 1.8.1. We say that $\overline{r}$ is *automorphic* if there exists an algebraic ℓ-adic representation $r\colon G_F \to \mathrm{GL}_n(\overline{\mathbf{Q}}_\ell)$ which is automorphic, whose reduction is $\overline{r}$.

Automorphy lifting theorems, which are the subject of this course, are of the form: given an ℓ-adic representation r such that $\overline{r}$ is automorphic, plus hypotheses, conclude that r is automorphic.

Potential automorphy lifting theorems are of the form: given an ℓ-adic representation r such that $\overline{r}$ is automorphic, plus hypotheses, conclude that there exists a finite Galois extension E/F such that $r|_{G_E}$ is automorphic.

In practice, many desirable consequences of automorphy are already implied by potential automorphy.

2. AN AUTOMORPHY LIFTING THEOREM FOR TOTALLY REAL FIELDS

2.1. **Main Theorem.** We begin by stating the main theorem that we will spend the next few weeks working to prove.

Theorem 2.1.1. *Assume $\ell > 2$, and fix $\iota: \overline{\mathbf{Q}}_\ell \xrightarrow{\sim} \mathbf{C}$. Let F be a totally real field, and $r: G_F \to \mathrm{GL}_2(\overline{\mathbf{Q}}_\ell)$ is a regular, algebraic ℓ-adic representation. Suppose there exists a regular algebraic cuspidal automorphic representation π of $\mathrm{GL}_2(\mathbf{A}_F)$ such that*

- $\overline{r_{\ell,\iota}(\pi)} \cong \overline{r}$,
- $\mathrm{HT}_\tau(r_{\ell,\iota}(\pi)) = \mathrm{HT}_\tau(r)$ *for all* $\tau: F \hookrightarrow \overline{\mathbf{Q}}_\ell$.

Assume furthermore:

1. $\overline{r}|_{G_{F(\zeta_\ell)}}$ *is irreducible.*
2. *ℓ is unramified in F and r is crystalline at all $v \mid \ell$ (equivalently, $\mathrm{WD}(r|_{G_{F_v}})$ is unramified for all $v \mid \ell$).*
3. *For all $v \mid \ell$, π_v is unramified (which is equivalent to $\mathrm{rec}(\pi_v)$ being unramified).*
4. *There exists a such that $\mathrm{HT}_\tau(r) \subset \{a, a+1, a+2, \dots, a+\ell-2\}$ for all $\tau: F \hookrightarrow \overline{\mathbf{Q}}_\ell$ ("Fontaine-Laffaille case").*

Then r is automorphic.

2.2. **Modularity of elliptic curves.** Before moving on to the proof, we give an application.

Corollary 2.2.1. *Suppose E/F is an elliptic curve over a totally real field F.*

1. *There exists a finite, totally real Galois extenson F'/F such that $E \times_F F'$ is automorphic.*
2. *$L(E, s)$ (defined for $\mathrm{Re}\, s > 3/2$) has meromorphic continuation to $\mathbf{C}$, and satisfies a functional equation*
$$\Lambda(E, s) = \pm N^{1-s} \Lambda(E, 2-s),$$
where $\Lambda(E, s) = L(E, s)(2\pi^{-s}\Gamma(s))^{[F:\mathbf{Q}]}$, and N is the conductor of E.

To E and a place v of F, there is a Weil-Deligne representation $\mathrm{WD}(E \times_F F_v)$, which is defined over $\overline{\mathbf{Q}}$ (and even "morally defined over $\mathbf{Q}$"). Consider the ℓ-adic G_F-representation $H^1(E \times_F \overline{F}; \mathbf{Q}_\ell)$. This is dual (as a G_F-representation) to the Tate module $V_\ell(E)$.

Proposition 2.2.2 (Fontaine)**.** *There is a 2-dimensional (Frobenius-semisimple) Weil-Deligne representation $\mathrm{WD}(E/F_v)$ defined over $\overline{\mathbf{Q}}$, such that for any ℓ and any embedding $j: \overline{\mathbf{Q}} \hookrightarrow \overline{\mathbf{Q}}_\ell$,*
$$j(\mathrm{WD}(E/F_v)) \cong \mathrm{WD}(H^1_{\text{ét}}(E \times_F \overline{F}; \mathbf{Q}_\ell)|_{G_{F_v}}).$$

You can think of this as saying that the ℓ-adic representations are "independent of ℓ". Conjecturally, an analogue holds for any motive.

Example 2.2.3. This representation captures geometric properties of E. For example,

- E has good reduction at v if and only if $\mathrm{WD}(E/F_v)$ is unramified.
- E has multiplicative reduction at v if and only if $\mathrm{WD}(E/F_v) \cong (\rho, N)$ where ρ is unramified and $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

- E has potentially good reduction at v if and only if $N=0$.

Definition 2.2.4. We say that E is *automorphic* if there exists an algebraic cuspidal automorphic representation π such that:

(1) $\mathrm{WD}(E/F_v)\cong \mathrm{rec}(\pi_v)$ for all v,
(2) $\mathrm{HC}_\tau(\pi_\infty)=\{0,-1\}$ for all $\tau: F\hookrightarrow \mathbf{C}$.

By Proposition 2.2.2, this is equivalent to the automorphy of $V_\ell(E)^\vee$ for one ℓ, and to the automorphy for all ℓ. We define the L-function of E to be

$$L(E,s)=\prod_v \det(1-\#k(v)^{-s}\,\mathrm{Frob}_v)|^{-1}_{\mathrm{WD}(E/F_v)^{I_{F_v}},N=0}.$$

For almost all v (namely the unramified ones), the local factor is

$$\left(1-\mathrm{Tr}(\mathrm{Frob}_v)\cdot \#k(v)^s+\#k(v)^{1-2s}\right)^{-1}.$$

Furthermore,

$$\#E(k(v))=1+\#k(v)-\mathrm{Tr}(\mathrm{Frob}_v).$$

Note that we need to take fixed points for I_{F_v} to get an action of $\mathrm{Frob}_v\in W_{F_v}/I_{F_v}$.

Remark 2.2.5. There are other ways to write down the Euler factors at "good" places, but the only way I know to define the "bad" factors is via Galois representations.

2.3. **Proof of Corollary 2.2.1.** We'll first show how to deduce Corollary 2.2.1 from Theorem 2.1.1.

2.3.1. *Initial reduction.* We can assume that E doesn't have CM over $\overline{F}$, i.e. $\mathrm{End}(E/\overline{F})=\mathbf{Z}$. This is because the CM case is easy.

2.3.2. *Auxiliary elliptic curve.* Next choose $E_0/\mathbf{Q}$ a CM elliptic curve, which has CM by an imaginary quadratic field M. There exists an algebraic Grossencharacter $\chi:\mathbf{A}_M^\times/M^\times\to\mathbf{C}^\times$ such that

$$V_\ell(E_0)^\vee\cong \mathrm{Ind}_M^{G_\mathbf{Q}}\, r_{\ell,\iota}(\chi).$$

For some choice of $\tau: M\hookrightarrow\mathbf{C}$, it will be the case that $\mathrm{HC}_\tau(\chi)=\{-1\}$ and $\mathrm{HC}_{\overline{\tau}}(\chi)=\{0\}$.

2.3.3. *Modular curves.* Now choose odd primes ℓ_1,ℓ_2 such that:

- E_0, E have good reduction at the ℓ_i,
- ℓ_1,ℓ_2 are unramified in F,
- $E[\ell_2]|_{G_{F(\zeta_2)}}$ is irreducible. (This is satisfied for almost all ℓ_2, by Serre's theorem that $G_F\twoheadrightarrow \mathrm{Aut}(E[\ell])$ for almost all ℓ, since E doesn't have CM.)
- $E_0[\ell_1]|_{G_{F(\zeta_{\ell_1})}}$ is irreducible. (This should be true for all but finitely many ℓ_1; it's at least easy to see that it's true for all ℓ_1 inert in M.)

Let X/F be the moduli space for elliptic curves A together with isomorphisms (as group schemes over F, i.e. as G_F-modules)

$$j_1: E_0[\ell_1]\xrightarrow{\sim} A[\ell_1],$$
$$j_2: E[\ell_2]\xrightarrow{\sim} A[\ell_2].$$

Remark 2.3.1. After base changing to $\mathbf{C}$, the associated complex-analytic variety is $\Gamma(\ell_1\ell_2)\backslash\mathbf{H}$. Hence X is geometrically irreducible. I don't know a proof of this fact that doesn't use our knowledge of the associated complex-analytic variety.

2.3.4. *Points over local fields.* We are going to define some local points on X.

- If $v \mid \ell_1$, choose a finite unramified extension L'_v/F_v such that $G_{L'_v}$ acts trivially on $E[\ell_2]$ and $E_0[\ell_2]$. Then E_0 gives a point on $X(L'_v)$, since we can take $j_1 = \mathrm{Id}\colon E_0[\ell_1] = E_0[\ell_1]$ and $j_2\colon E[\ell_2] \xrightarrow{\sim} E_0[\ell_2]$.
- Similarly, if $v \mid \ell_2$ we choose a finite unramified extension L'_v/F_v such that $G_{L'_v}$ acts trivially on $E[\ell_1]$ and $E_0[\ell_1]$. Then E gives a point in $X(L'_v)$.

Let $\Omega_v \subset X(L'_v)$ be the locus of (A, j_1, j_2) where A has good reduction. (Note that the triviality of the Galois action on the torsion implies that the reduction is already semistable – no additive reduction – so the good reduction property is exactly detected by integrality of the j-invariant.)

2.3.5. *Points over global fields.* We now state a result that tells us that we can find a rational point over a global field satisfying certain desired local conditions.

Proposition 2.3.2 (Moret-Bailly)**.** *Let K^{avoid}/K be a finite Galois extension of number fields. Let S be a finite set of places of K. Let S be a finite set of places of K. If $v \in S$, let L'_v/K_v be a finite Galois extension.*

Let X/K be a smooth and geometrically connected variety. Suppose that for all $v \in S$, there exists a non-empty subset $\Omega_v \subset X(L'_v)$ which is $\mathrm{Gal}(L'_v/K_v)$-invariant. Then: there exists L/K which is finite and Galois, and linearly disjoint from K^{avoid}, a point $p \in X(L)$, and isomorphisms $L_w \cong L'_v$ for all $w \mid v \in S$, such that $p \in \Omega_v \subset X(L_w) \cong X(L'_v)$ for all $w \mid v \in S$.

We apply this with:

- $K = F$,
- $S = \{v \mid \infty\ell_1\ell_2\}$,
- $L'_v = L'_v$ if $v \mid \ell_1\ell_2$,
- $L'_v = \mathbf{R}$ if $v \mid \infty$. (Note that this guarantees that L is also totally real.)
- Ω_v as above. For $v \nmid \infty$, we know that Ω_v is non-empty because we've designed it to contain (E_0, j_1, j_2) or (E_1, j_1, j_2). For $w \mid \infty$, we need to know that $X(\mathbf{R}) \neq$. In the $\mathbf{R}$ case, we need to match the complex conjugations. The point here is that complex conjugation is an element of $\mathrm{GL}_2(\mathbf{F}_{\ell_1})$ with order 2 and determinant -1, and there is only one conjugacy class of such elements. Hence a point is given by, for example, taking $E = E_0$, $j_1 = \mathrm{Id}$, and j_2 to be an isomorphism which necessarily exists by the preceding considerations.
- $K^{\mathrm{avoid}} = F(E[\ell_1\ell_2], E_0[\ell_1\ell_2], \zeta_{\ell_1\ell_2})$. The point of this is to preserve the irreducibility hypothesis on the residual Galois representation, and that for a field linearly disjoint from K^{avoid}, the Galois action doesn't change.

So we get that there exists a finite Galois and totally real extension L/K, and an elliptic curve A/L, with isomorphisms

$$j_1\colon E_0[\ell_1] \xrightarrow{\sim} A[\ell_1],$$
$$j_2\colon E[\ell_2] \xrightarrow{\sim} A[\ell_2].$$

such that A has good reduction above $\ell_1\ell_2$. In particular, $A[\ell_1]$ is irreducible as a representation of $G_{L(\zeta_{\ell_1})}$ and $A[\ell_2]$ is irreducible as a representation of $G_{L(\zeta_{\ell_2})}$.

2.3.6. *Base change for grossencharacters.* We are then going to use Theorem 2.1.1 to deduce that A is automorphic over L from the fact that E_0 is automorphic over L (using j_1), and then that E is automorphic over L from the fact that A is automorphic over L (using j_2).

We explain this carefully. Recall that since E_0 has CM by the imaginary quadratic field M, we can use CM theory to show that there exists a grossencharacter $\chi\colon \mathbf{A}_M^\times/M^\times \to \mathbf{C}^\times$ and an embedding $\tau\colon M \hookrightarrow \mathbf{C}$ such that $\mathrm{HC}_\tau(\chi_\infty) = \{-1\}$, $\mathrm{HC}_{\overline{\tau}}(\chi_\infty) = \{0\}$, and

$$\mathrm{Ind}_{G_M}^{G_\mathbf{Q}} r_{\iota,\ell}(\chi) \simeq (V_\ell E_0)^\vee.$$

Now, this tells us that:

$$(V_\ell E_0)^\vee|_{G_L} \simeq \mathrm{Ind}_{G_{ML}}^{G_L}(r_{\iota,\ell}(\chi)|_{G_{ML}}) = \mathrm{Ind}_{G_{ML}}^{G_L} r_{\iota,\ell}(\chi\circ N_{ML/M}).$$

The second equality is by compatibility of the Artin map with the norms. Furthermore, since ML/L is certainly solvable, the base change theorems we discussed in §1.7.5 imply that $\mathrm{Ind}_{G_{ML}}^{G_L} r_{\iota,\ell}(\chi\circ N_{ML/M})$ is automorphic, i.e.

$$\mathrm{Ind}_{G_{ML}}^{G_L} r_{\iota,\ell}(\chi\circ N_{ML/M}) = r_{\iota,\ell}(\pi_0)$$

for some regular algebraic cuspidal automorphic representation π_0 of $\mathrm{GL}_n(\mathbf{A}_L)$ such that $\mathrm{HC}_\tau(\pi_{0,\infty}) = \{0,-1\}$ for all $\tau\colon L\hookrightarrow\mathbf{C}$.

2.3.7. *Transferring automorphy.* Now, we apply Theorem 2.1.1 to the representation $r = (V_{\ell_1}A)^\vee$, using that

$$\overline{r} = (A[\ell_1])^\vee \simeq E_0[\ell_1]^\vee \simeq \overline{r_{\iota,\ell_1}(\pi_0)},$$

so we have automorphy mod ℓ. We check that the various conditions in the theorem are satisfied by construction of A. This gives us a regular algebraic cuspidal automorphic representation π_1 of $\mathrm{GL}_2(\mathbf{A}_L)$ such that $r_{\iota,\ell_1}(\pi_1) \simeq (V_{\ell_1}A)^\vee$. We then proceed similarly with $(V_{\ell_2}A)^\vee$.

2.3.8. We record the following:

Proposition 2.3.3. *If r, r' are two semi-simple algebraic ℓ-adic representations of G_F, then the following are equivalent:*

(i) r *is isomorphic to* r'.
(ii) $\mathrm{WD}(r|_{G_{F_v}}) \simeq \mathrm{WD}(r'|_{G_{F_v}})$ *for all* v.
(iii) $\mathrm{WD}(r|_{G_{F_v}}) \simeq \mathrm{WD}(r'|_{G_{F_v}})$ *for all but finitely many* v.

Proof. We just need to show that (iii) $\Longrightarrow$ (i). This hypothesis implies that for almost all v, $r|_{G_{F_v}}$ is unramified and we have:

$$\operatorname{tr} r(\mathrm{Frob}_v) = \operatorname{tr}[\mathrm{WD}(r|_{G_{F_v}})(\mathrm{Frob}_v)] = \operatorname{tr}[\mathrm{WD}(r'|_{G_{F_v}})(\mathrm{Frob}_v)] = \operatorname{tr} r'(\mathrm{Frob}_v).$$

Now, the *Cebotarev density theorem* tells us that if F_S is the maximal extension of F unramified outside of S and G_S is its Galois group, then the set of Frobenius elements Frob_v for $v \notin S$ is dense in G_S. Hence $\operatorname{tr} r = \operatorname{tr} r'$, and then $r \simeq r'$ by semisimplicity. □

2.3.9. *Application to L-functions.* We now explore an important consequences of Corollary 2.2.1.

Corollary 2.3.4. *The L-function $L(E,s)$ (defined for $\Re s > 3/2$) has a meromorphic continuation to $\mathbf{C}$ and satisfies a functional equation:*

$$\mathscr{L}(E,s) = \epsilon N^{1-s} \mathscr{L}(E, 2-s)$$

Here, $\mathscr{L}(E,s) = L(E,s)(2\pi^{-s}\Gamma(s))^{[F:\mathbf{Q}]}$, $N \in \mathbf{Z}_{>0}$ is the conductor of E, and $\epsilon = \pm 1$.

Proof. Consider the fields $F \subseteq F_i \subseteq L$ with L/F_i solvable. There exists a regular algebraic cuspidal automorphic representation π_i of $\mathrm{GL}_2(\mathbf{A}_{F_i})$ such that $\mathrm{rec}(\pi_{i,v}) \simeq \mathrm{WD}(E/F_{i,v})$ for all v and $\mathrm{HC}_\tau(\pi_{i,\infty}) = \{0,-1\}$ for all τ by the solvable base change theorems. In other words, $r_\ell(\pi_i) = (V_\ell E)^\vee|_{G_{F_i}}$.

As a consequence, (by the definition of the L-functions on each side), we have

$$L(\pi_i, s) = L(E_{F_i}, s) = L((V_\ell E)^\vee|_{G_{F_i}}, s).$$

2.3.10. *Brauer induction.* Now, a theorem of Brauer tells us that there exists $n_i \in \mathbf{Z}$ and $\chi_i\colon \mathrm{Gal}(L/F_i) \to \overline{\mathbf{Q}}^\times$ such that

$$\sum_i n_i \operatorname{Ind}_{\mathrm{Gal}(L/F_i)}^{\mathrm{Gal}(L/F)} \chi_i = 1 \tag{2.3.1}$$

in the Grothendieck ring of representations of the finite group $\mathrm{Gal}(L/F)$, where the right hand side of (2.3.1) is the trivial representation. We can replace the finite group $\mathrm{Gal}(L/F)$ with G_F and $\mathrm{Gal}(L/F_i)$ with G_{F_i}, and the statement is still true. This implies that:

$$\begin{aligned}(V_\ell E)^\vee &\simeq \sum_i n_i \operatorname{Ind}_{G_{F_i}}^{G_F} [\chi_i \otimes ((V_\ell E)^\vee|_{G_{F_i}})] \\ &\simeq \sum_i n_i \operatorname{Ind}_{G_{F_i}}^{G_F} (\chi_i \otimes r_\ell(\pi_i)) \\ &\simeq \sum_i n_i \operatorname{Ind}_{G_{F_i}}^{G_F} (r_\ell(\pi_i')).\end{aligned}$$

Now, we have $r_\ell(\chi) \otimes r_\ell(\pi) \simeq r_\ell(\pi \otimes (\chi \circ \det)) =: r_\ell(\pi_i')$.

We can appeal to the following (easy) properties of L-functions:

(i) $L(r \oplus r', s) = L(r,s)L(r',s)$.
(ii) $L(\operatorname{Ind}_{G_{F'}}^{G_F} r, s) = L(r,s)$.

These imply that:

$$L(E,s) = L((V_\ell E)^\vee, s) = \prod_i L(\pi_i', s)^{n_i}.$$

2.3.11. Now, we have that $\mathscr{L}(\pi_i, s) = (2\pi^{-s}\Gamma(s))^{[F_i:\mathbf{Q}]}L(\pi_i, s)$. Taking dimensions in (2.3.1), we see that $1 = \sum_i n_i[F_i : F]$ and therefore that $[F : \mathbf{Q}] = \sum_i n_i[F_i : \mathbf{Q}]$. This implies that we also have:

$$\mathscr{L}(E,s) = \mathscr{L}((V_\ell E)^\vee, s) = \prod_i \mathscr{L}(\pi_i', s)^{n_i}.$$

Now, the L-function of an automorphic representation may be written as a certain integral. This can be used to show that $L(\pi_i, s)$ has analytic continuation to $\mathbf{C}$ with a functional equation:

$$\mathscr{L}(\pi_i, s) = \epsilon N_i^{1-s}\mathscr{L}(\pi_i, 2-s),$$

for $\epsilon = \pm 1$ and $N_i \in \mathbf{Z}_{>0}$. This implies that $L(E,s)$ has meromorphic continuation (no longer necessarily analytic, since the n_i could very well be negative) to all of $\mathbf{C}$, and multiplying the functional equations tells us that:

$$\mathscr{L}(E,s) = \left(\prod_i \epsilon^{n_i}\right)\left(\prod_i N_i^{n_i}\right)^{1-s}\mathscr{L}(E, 2-s).$$

□

2.4. **Initial Reductions.** We begin with some reductions towards the proof of Theorem 2.1.1.

2.4.1. We can assume that $a = 0$ by replacing r with $r \otimes \epsilon_\ell^a$ and π by $\pi \otimes |\det|^a$. (In our convention, the cyclotomic character has Hodge-Tate weight -1.)

2.4.2. We get rid of ramification. More precisely, we will reduce to the case where:

- for all v, $\mathrm{WD}(r|_{G_v})^{ss}$ is unramified, i.e. the Weil-Deligne representation (ρ, N) is unramified.
- For all v, $\mathrm{rec}(\pi_v)^{ss}$ is unramified. (By what we know about the Local Langlands correspondence, this is equivalent to π_v having a non-zero fixed vector under the Iwahori subgroup. This is called the "semistable case".)
- We may assume that $[F : \mathbf{Q}]$ is even.
- We may assume that if π_v is ramified, then $\ell \mid \# k(v)^\times$.

To get these reductions, we use solvable base change (§1.7.5) and the following lemma.

Lemma 2.4.1. *Let S be a finite set of places of a number field K. For each $v \in S$ let L_v'/K_v be a finite Galois extension. Then there is a finite solvable Galois extension L/K such that if $w \mid v \in S$, then $L_w \cong L_v'$ as a K_v-algebra. Moreover, if K^{avoid}/K is any finite extension then we can choose L to be linearly disjoint from K^{avoid}.*

This is a consequence of class field theory; we will omit the proof.

Remark 2.4.2. The statement of Proposition 2.3.2 looks similar, but there are a couple of major differences. Here we are not asking for existence of rational points. On the other hand, we *are* demanding that the global extension be solvable.

To get the claimed reductions, we look at the finitely many places where $\mathrm{WD}(r|_{G_{F_v}})^{\mathrm{ss}}$ is ramified. Since the image of inertia is finite, we can make a finite local extension killing it. The same argument applies for $\mathrm{rec}(\pi_v)^{\mathrm{ss}}$. The even-ness of the global degree is easy to arrange – we can just make all the local extensions have even degree. Finally, we can make unramified local extensions to get the cardinality of the residue field to be as desired.

The hypotheses of the theorem will still obviously be satisfied, except for the condition that $r|_{G_{F(\zeta_\ell)}}$ is irreducible. To arrange this, we take K^{avoid} to be $\overline{F}^{\ker \overline{r}}(\zeta_\ell)$. Since $\overline{r}$ factors through $\mathrm{Gal}(K^{\mathrm{avoid}}/F)$, the linear disjointness in the diagram

$$\begin{tikzcd} & LK^{\mathrm{avoid}} \arrow[dl, dash] \arrow[dr, dash] & \\ L \arrow[dr, dash] & & K^{\mathrm{avoid}} \arrow[dl, dash] \\ & K & \end{tikzcd}$$

implies that $\mathrm{Gal}(LK^{\mathrm{avoid}}/L) \cong \mathrm{Gal}(K^{\mathrm{avoid}}/K)$, hence the image of $\overline{r}|_{G_{L(\zeta_\ell)}}$ is unchanged.

2.4.3. We will show that we may assume $\det r = \det r_\ell(\pi)$. Indeed,

$$\mathrm{HT}_\tau(\det r) = \sum_{h \in \mathrm{HT}_\tau(r)} h = \sum_{h \in \mathrm{HT}_\tau(r_\ell(\pi))} h = \mathrm{HT}_\tau(r_\ell(\pi)).$$

Hence $\psi = (\det r)(\det r_\ell(\pi))^{-1}$ is a character with $\mathrm{HT}_\tau(\psi) = \{0\}$ and $\psi \equiv 1 \pmod{\mathfrak{m}_{\overline{\mathbf{Q}}_\ell}}$, so that ψ has finite image. The reason is that the corresponding Grossencharacter

$$\psi_0 \circ \mathrm{Art}_F : \mathbf{F}^\times \backslash \mathbf{A}_F^\times / (F_\infty^0)^\times U \to \mathbf{C}^\times$$

is U_ℓ-invariant for an open subgroup of $\mathbf{Z}_p^\times$ (by the fact that the Hodge-Tate weight is 0), so the double coset space only has finite order.

Since $\psi \equiv 1 \pmod{\mathfrak{m}_{\overline{\mathbf{Q}}_\ell}}$, then ψ has ℓ-power order, say ℓ^a. Furthermore, since $\mathrm{WD}(r|_{G_{F_v}})$ and π_v were unramified, we even know that ψ is crystalline. (Determinants of crystalline representations are crystalline, and tensor products of crystalline are crystalline.) Therefore, it is unramified (think to our explicit construction of the WD representation for a Grossencharacter). Then $\phi = \psi^{\frac{1-\ell^a}{2}}$ has the properties that ϕ has finite image (hence is crystalline with Hodge-Tate weight 0), $\phi \equiv 1 \pmod{\mathfrak{m}_{\overline{\mathbf{Q}}_\ell}}$, and $\phi^2 = \psi$.

Replacing r by $r \otimes \phi^{-1}$ doesn't change the mod ℓ representation, or any of the hypotheses, so the theorem applied to it would show that it is automorphic. But they are the same after a finite abelian base change trivializing the finite-order character ϕ. Hence, by solvable base change (§1.7.5), the question of automorphy is equivalent for r and $r \otimes \phi^{-1}$.

2.4.4. Finally, we reduce to showing that π_v is unramified *everywhere*. To arrange this, we will follow the steps:

(i) Show that we may assume $\mathrm{rec}(\pi_v)$ has $N = 0$ everywhere.

(ii) We then repeat the argument of §2.4.2 to ensure that $\mathrm{rec}(\pi_v)$ is unramified everywhere. The point is that performing a base change cannot kill N, but if we already know that $N = 0$, then a solvable base change can kill off the ramification.

Why not skip §2.4.2 and directly do §2.4.4 (whose conclusion is stronger)? The reduction (i) will actually be quite non-trivial, and for it we'll need the earlier steps.

We will first explain (i). Something new is happening here, because we can't just make a base change (N doesn't change under base change). For the first time, we will need to use the theory of congruences between automorphic forms. Since automorphic forms are analytic objects, to get started with congruences we will need an algebraic theory of automorphic forms, and this will require a big digression.

3. Automorphic forms for quaternion algebras

3.1. **The Jacquet-Langlands correspondence.** Let D be the quaternion algebra with center F, ramified exactly at all ∞ places (this exists since we have arranged $[F:\mathbf{Q}]$ to be even). This means that for $v \mid \infty$, $D_v \cong \mathbf{H}$ and for $v \nmid \infty$, $D_v \cong M_{2\times 2}(F_v)$.

We will define a space of automorphic forms on $D^\times$,

$$\mathscr{A}_0(D^\times \backslash (D \otimes_F \mathbf{A}_F)^\times, \chi)$$

where $\chi \colon \mathbf{A}_F^\times / F^\times \to \mathbf{C}^\times$. This is the set of functions $\varphi \colon D^\times \backslash (D \otimes_F \mathbf{A}_F)^\times \to \mathbf{C}^\times$ with the same conditions as before: smooth, $D_\infty^\times$-finite (this is simpler because the group is compact mod center), no growth conditions because the space is compact, and

$$\varphi(gz) = \chi(z)\varphi(g) \text{ for all } z \in \mathbf{A}_F^\times.$$

Using the trace formula, Jacquet-Langlands proved:

Theorem 3.1.1 (Jacquet-Langlands)**.** *There is a decomposition*

$$\mathscr{A}_0(D^\times \backslash (D \otimes_F \mathbf{A}_F)^\times, \chi) \cong \bigoplus \pi^D$$

where π^D is an irreducible representation of $(D \otimes \mathbf{A}_F)^\times$, fitting into one of the following two cases:

- $\pi^D = \phi \circ \det$ *(where* $\det$ *is what is usually called the "reduced norm") where ϕ is a Grossencharacter $\mathbf{F}^\times \backslash \mathbf{A}_F^\times \to \mathbf{C}^\times$ with $\phi^2 = \chi$, or*
- $\pi^D \cong \pi^\infty \otimes \pi_\infty^D$ *where π is a cuspidal automorphic representation of* $\mathrm{GL}_2(\mathbf{A}_F)$ *such that* $\mathrm{HC}_\tau(\pi_\infty) = \{1 + n_\tau + s_\tau, s_\tau\}$ *with $n_\tau \in \mathbf{Z}_{\geq 0}$ and $s_\tau \in \mathbf{C}$, and*

$$\pi_\infty^D = \bigotimes_{\tau \colon F \hookrightarrow \mathbf{R}} \left(\mathrm{Sym}^{n_\tau}(\mathbf{C}^2) \otimes (\det)^{s_\tau + 1/2}\right)^\vee$$

and any such π occurs as long as $\chi_{\pi^\infty} = \chi$ with $\mathrm{HC}_\tau(\chi) = \{1 + n_\tau + 2s_\tau\}$.

3.2. **A model for modular forms.** We are now going to look at a specific space of modular forms inside $\mathscr{A}_0(D^\times \backslash (D \otimes_F \mathbf{A}_F)^\times, \chi)$.

3.2.1. Let

$$k := \{(n_\tau, m_\tau) \mid \tau\colon F \hookrightarrow \mathbf{R}\,;\, n_\tau \in \mathbf{Z}_{\geq 0}, m_\tau \in \mathbf{Z}\}$$

for some choice of m_τ, n_τ such that $w := 1 + n_\tau + 2m_\tau$ is independent of τ. We think of k as the "weight" of the modular forms that we will eventually define.

Definition 3.2.1. Let $\chi\colon \mathbf{A}_F^\times/F^\times \to \mathbf{C}^\times$ be a continuous Grossencharacter, such that $\chi|_{(F_\infty^\times)^0} = \mathrm{Nm}_{F/\mathbf{Q}}^{1-w}$, i.e.

$$\chi(x) = \prod_{\tau\colon F \hookrightarrow \mathbf{R}} (\tau x)^{1-w}$$

for $x \in (F_\infty^\times)^0$. (Unless w is independent of τ, there aren't going to be any modular forms.) Set

$$S_{k,\chi}(\mathbf{C}) = \mathrm{Hom}_{D_\infty^\times}\Big(\bigotimes_{\tau\colon F\hookrightarrow \mathbf{R}} (\mathrm{Sym}^{n_\tau}(\mathbf{C}^2) \otimes (\det)^{m_\tau})^\vee, \mathscr{A}_0(D^\times \backslash (D \otimes \mathbf{A})^\times, \chi)\Big).$$

This has an action of $\mathrm{GL}_2(\mathbf{A}_F^\infty) \cong (D \otimes \mathbf{A}_F^\infty)^\times$.

3.2.2. *The Jacquet-Langlands correspondence.* Jacquet-Langlands gave a description of the space $S_{k,\chi}(\mathbf{C})$ (deduced from Theorem 3.1.1). There are two pieces.

(1) The first piece is a direct sum over π, regular cuspidal aglebraic representations of $\mathrm{GL}_2(\mathbf{A}_F)$, of

$$\pi^\infty \otimes ||\det||^{m_\tau}.$$

(In the notation of Theorem 3.1.1, $s_\tau + 1/2$ is what we're now calling the integer m_τ. So s_τ should be a half integer.)

The π appearing are constrained by the conditions that

$$\mathrm{HC}_\tau(\pi_\infty) = \{n_\tau + m_\tau + 1, m_\tau\}$$

and

$$\chi_\pi = \chi ||\cdot||^{-1}.$$

Remark 3.2.2. For the representation $\mathrm{Sym}^n(\mathbf{C}^2)$ of $\mathrm{GL}_2(\mathbf{C})$, the Harish-Chandra parameter is $\{(n+1/2, -1/2)\}$, while the highest weight is $(n, 0)$. So you see a shift here. The shifts appearing in the theory are not just an artifact of poor presentation, but a fundamental fact of life.

(2) The second piece is a direct sum of $\phi^\infty \circ \det$, for $\phi\colon \mathbf{A}_F^\times/F^\times \to \mathbf{C}^\times$ with $\phi^2 = \chi$, if $n_\tau = 0$ for all τ.

3.2.3. For a finite open subgroup $\mathrm{GL}_2(\mathbf{A}_F^\infty)$, we will write $S_{k,\chi}(U, \mathbf{C}) := S_{k,\chi}(\mathbf{C})^U$. We claimed that we would be able to give an algebraic model for this, and we are going to justify that now.

Lemma 3.2.3. *We have an isomorphism of* $S_{k,\chi}(\mathbf{C})$ *with the space of functions*

$$\varphi\colon D^\times \backslash (D \otimes \mathbf{A}_F)^\times \to \bigotimes_\tau \mathrm{Sym}^{n_\tau}(\mathbf{C}^2) \otimes (\det)^{m_\tau}$$

such that

- $\varphi(gh) = h^{-1}\varphi(g)$ *for* $h \in D_\infty^\times$,
- $\varphi(gz) = \chi(z)\varphi(g)$,
- φ *is invariant under right translation by some open compact subgroup of* $\mathrm{GL}_2(\mathbf{A}_F^\infty)$.

Proof. Let's write $S_{k,\chi}(\mathbf{C})'$ for the space in question in Lemma 3.2.3. Let's show how to go from $\varphi \in S_{k,\chi}(\mathbf{C})'$ to $F \in S_{k,\chi}(\mathbf{C})$.

Looking at the definition of $S_{k,\chi}(\mathbf{C})$, from φ we need to build a function that takes in $\lambda \in (\bigotimes_\tau \operatorname{Sym}^{n_\tau}(\mathbf{C}^2) \otimes (\det)^{m_\tau})^\vee$ and $g \in (D \otimes \mathbf{A}_F)^\times$ and spits out a number. Well, for any such g we get $\varphi(g) \in \bigotimes_\tau \operatorname{Sym}^{n_\tau}(\mathbf{C}^2) \otimes (\det)^{m_\tau}$. Then we set

$$F(\lambda)(g) = \lambda(\varphi(g)).$$

The inverse is easy to write down. Then you check that the conditions match up, i.e. the $D_\infty^\times$-linearity on one side translates into the first condition. $\square$

3.2.4. We now rewrite $S_{k,\chi}(\mathbf{C})'$ in a way that parallels the maniuplations in §1.7.1.

Lemma 3.2.4. *We have an isomorphism of $S_{k,\chi}(\mathbf{C})$ with the space of functions*

$$f : (D \otimes \mathbf{A}_F^\infty)^\times \to \bigotimes_\tau \operatorname{Sym}^{n_\tau}(\mathbf{C}^2) \otimes (\det)^{m_\tau}$$

such that

- $f(\delta g) = \delta f(g)$ *for all* $\delta \in D^\times$,
- $f(gz) = \widetilde{\chi}(z) f(g)$ *for all* $z \in (\mathbf{A}_F^\infty)^\times$,
- f *is invariant under translation by some open compact subgroup.*

Here $\widetilde{\chi}$ is as in (1.7.1)*; it is the same operation in taking an algebraic Grossencharacter to $\widetilde{\chi} : \mathbf{A}_F^\times \to \mathbf{C}^\times$ that factored through the infinite component.*

Proof. Let $S_{k,\chi}(\mathbf{C})''$ be the space of functions in question. We will compare it to the $S_{k,\chi}(\mathbf{C})'$ model from Lemma 3.2.3.

Given $\varphi \in S_{k,\chi}(\mathbf{C})'$, we make f by sending $g \in (D \otimes \mathbf{A}_F)^\times$

$$f(g) := g_\infty \varphi(g)$$

Is it well-defined? For $h \in D^\times$, $gh \mapsto g_\infty h \varphi(gh) = g_\infty \varphi(g)$.

In the other direction, we send f to $(g \mapsto g_\infty^{-1} f(g^\infty))$.

$\square$

The latter description works with $\overline{\mathbf{Q}}$ or $\overline{\mathbf{Q}}_\ell$ instead of $\mathbf{C}$ (since we have purged appearances of $D_\infty^\times$). This gives a way to make sense of $S_{k,\chi}(\overline{\mathbf{Q}}_\ell)$ or $S_{k,\chi}(\overline{\mathbf{Q}})$, which are "forms" of $S_{k,\chi}(\mathbf{C})$, in the sense that after tensoring up to $\mathbf{C}$ via $\overline{\mathbf{Q}} \hookrightarrow \mathbf{C}$ or $\overline{\mathbf{Q}}_\ell \xrightarrow{\sim} \mathbf{C}$ they recover $S_{k,\chi}(\mathbf{C})$.

3.2.5. Now we basically undo the condition from before to get the ℓ-adic version.

Lemma 3.2.5. *We have an isomorphism of $S_{k,\chi}(\overline{\mathbf{Q}}_\ell)$ with the space of functions*

$$\varphi : D^\times \backslash (D \otimes \mathbf{A}_F^\infty)^\times \to \bigotimes_{\tau : F \hookrightarrow \overline{\mathbf{Q}}_\ell} \operatorname{Sym}^n(\overline{\mathbf{Q}}_\ell^2) \otimes (\det)^{m_\tau}$$

satisfying the conditions:

- $\varphi(gz) = \chi^{(\ell)}(z)\varphi(g)$ *for all* $z \in (\mathbf{A}_F^\infty)^\times$*, where $\chi^{(\ell)}$ is the ℓ-adic character associated to χ in* (1.7.2).
- *There exists an open compact subgroup* $U \subset \mathrm{GL}_2(\mathbf{A}_F^\infty)$ *such that* $\varphi(gu) = u_\ell^{-1}\varphi(g)$ *for all* $u \in U$.

Proof. The bijection is defined by taking f to the function $\varphi(g) = g_\ell^{-1} f(g)$. The inverse takes φ to the function $g \mapsto g_\ell(\varphi(g))$. $\square$

3.3. **Integral models for automorphic forms.** Now we're in better position to talk about integrality of automorphic forms. (Before, it would have been problematic if δ had a non-integral part at ℓ.)

Since $\chi^{(\ell)}$ is continuous, it has finite image, so there exists a finite extension $L/\mathbf{Q}_\ell$ such that $\chi^{(\ell)} \colon (\mathbf{A}_F^\infty)^\times \to \mathcal{O}_L^\times = \mathcal{O}^\times$. We may also assume that $L \supset \tau(F)$ for all τ.

Suppose that $\mathrm{pr}_\ell(U) \subset \mathrm{GL}_2(\mathcal{O}_{F,\ell})$. (We can always guarantee this by passing to a subgroup of finite index.) Let A be an $\mathcal{O}$-algebra. Then we can define a notion of "A-valued automorphic forms for D", namely the space of functions

$$\varphi \colon D^\times \backslash (D \otimes \mathbf{A}_F^\infty)^\times \to \bigotimes_{\tau \colon F \hookrightarrow \overline{\mathbf{Q}}_\ell} \mathrm{Sym}^{n_\tau}(A) \otimes (\det)^{m_\tau}$$

such that

- $\varphi(gz) = \chi^{(\ell)}(z)\varphi(g)$ for all $z \in (\mathbf{A}_F^\infty)^\times$,
- There exists an open compact subgroup $U \subset \mathrm{GL}_2(\mathbf{A}_F^\infty)$ such that $\varphi(gu) = u_\ell^{-1}\varphi(g)$ for all $u \in U$.

3.3.1. What acts on $\mathrm{Sym}^{n_\tau}(A^2) \otimes (\det)^{m_\tau}$? This is slightly tricky, as we have to pay attention to denominators. More precisely, we need $\chi^{(\ell)}(z)$ and u_ℓ^{-1} to act on $\mathrm{Sym}^{n_\tau}(A) \otimes (\det)^{m_\tau}$. Answer:

- $\mathrm{GL}_2(\mathbf{A}_F^{\infty,\ell}) \times \mathrm{GL}_2(\mathcal{O}_{F,\ell})$ for any $\mathcal{O}$-algebra A.
- If A is an L-algebra, we also get an action of $\mathrm{GL}_2(F_\ell)$.
- Finally, we get an action of $\mathrm{GL}_2(F_\ell) \cap M_{2\times 2}(\mathcal{O}_{F,\ell})$ for all A if $m_\tau \geq 0$ for all τ. This satisfies

$$(h \cdot \varphi)(g) = h_\ell \varphi(gh).$$

Taking U-invariants, we can identify $S_{k,\chi}(\mathbf{A}_F)^U$ with the *finite-dimensional* space

$$\bigoplus_{g \in D^\times \backslash (D \otimes \mathbf{A}_F^\infty)^\times / U} \left(\bigotimes_{\tau \colon F \hookrightarrow \overline{\mathbf{Q}}_\ell} \mathrm{Sym}^{n_\tau}(A^2) \otimes (\det)^{m_\tau} \right)^{U \cdot (\mathbf{A}_F^\infty)^\times \cap g^{-1} D^\times g}$$

by sending

$$\varphi \in S_{k,\chi}(\mathbf{A}_F)^U \mapsto (\varphi(g) \colon g \in D^\times \backslash (D \otimes \mathbf{A}_F^\infty)^\times / U).$$

(Check that the image is fixed by the action of $uz \in U(\mathbf{A}_F^\infty)^\times$, which is via $\chi^{(\ell)}(z)^{-1} u_\ell$.) The point is that the double coset space $D^\times \backslash (D \otimes \mathbf{A}_F^\infty)^\times / U$ is just a *finite set.*

3.4. **Growth with level structures.** We now discuss the "growth" of the spaces of modular forms $S_{k,\chi}(A)^U$ with the level structure U.

3.4.1. *Sufficiently small arithmetic groups.* Define $\Delta_{g,U}$ by the exact sequence

$$0 \to F^\times \to U \cdot (\mathbf{A}_F^\infty)^\times \cap g^{-1} D^\times g \to \Delta_{g,U} \to 0.$$

So $\Delta_{g,U}$ is both compact and discrete, hence finite.

Definition 3.4.1. We will call U *sufficiently small* for ℓ if $\ell \nmid \#\Delta_{g,U}$ for all g.

Assume that U is sufficiently small and that A is noetherian. Then we have the following facts:

(1) $S_{k,\chi}(A)^U$ is a finite A-module, and if A is an integral domain then it's torsion-free.
(2) If B/A is flat, then $S_{k,\chi}(A)^U \otimes_A B \xrightarrow{\sim} S_{k,\chi}(B)^U$. (Indeed, the invariants are the kernel of a certain map, and flat extensions commute with taking kernels.)
(3) If U is sufficiently small for ℓ, then $S_{k,\chi}(A)^U$ is a projective A-module and

$$S_{k,\chi}(A)^U \otimes_A B \xrightarrow{\sim} S_{k,\chi}(B)^U \text{ for all } B/A.$$

To see this, use the idempotent $e = (\sum_{\delta \in \Delta_{g,u}} \delta)\#\Delta_{g,u}^{-1}$, which projects onto the U-fixed part. Hence the U-fixed part is a summand whenever $\#\Delta_{g,u}$ is invertible. Since $\mathrm{Sym}^2(A)$ is free, after tensoring with B it's still free, and the U-fixed part is a summand of a free module, hence projective.

Lemma 3.4.2. *Suppose either*

(1) $[F(\zeta_\ell):F] > 2$, *or*
(2) $[F(\zeta_\ell):F] = 2$ *and there exists* $v_0 \nmid \ell$ *such that* $(\mathrm{Tr}\, u_{v_0})^2 \equiv 4 \det u_{v_0}$ *for all* $u \in U$.

Then U is sufficiently small for ℓ.

Remark 3.4.3. The second condition is satisfied, for example, when $U = \mathrm{Iw}_{v_0}^1 \times U^{v_0}$, where $\mathrm{Iw}_{v_0}^1$ is

$$\left\{ g \in \mathrm{GL}_2(\mathcal{O}_{F,v_0}) \colon g \equiv \begin{pmatrix} a & b \\ & a \end{pmatrix} \pmod{v_0} \right\}.$$

It seems like a really technical condition, but sometimes it's really necessary to push things as far as you can. For example, in Wiles' proof of Fermat's Last Theorem he needed to work with $\ell = 2$, where the first criterion (1) really isn't enough.

Proof. Suppose for $\delta \in D^\times$,$g^{-1}\delta g$ maps to an element of order ℓ in $\Delta_{g,u}$. In other words, $\delta^\ell \in F^\times$.

We will use the following basic facts about quaternion algebras:

- There is an involution $*\colon D \to D$, for which $F = \{\delta \in D : \delta^* = \delta\}$.
- We have $\mathrm{Tr}(\delta) = \delta + \delta^*$ and $\det \delta = \delta\delta^*$.

Now consider $(\delta/\delta^*)^\ell = 1$. If $\delta/\delta^* = 1$ then $\delta \in F^\times$, which contradicts its order being ℓ in $\Delta_{g,U}$. Hence δ/δ^* is an ℓth root of unity. So D contains $F(\zeta_\ell)$, but a quaternion algebra can only contain quadratic field extensions of its center. This proves (1).

For (2), the same argument implies that $[F[\delta/\delta^*]:F] = 2$. In $\mathcal{O}_{F[\delta]=F[\delta/\delta^*]}$ (equality because both have degree 2 over F), we have $(\delta) = (\delta^*)$, as δ/δ^* is a root of unity. The assumption $v_0 \nmid \ell$ implies that v_0 is unramified in $F[\delta/\delta^*]$. So we can alter δ by an element of $F^\times$ to make it a unit at v_0 (using that v_0 is unramified and considering the split/inert cases). By assumption,

$$(\delta+\delta^*)^2 \equiv 4\delta\delta^* \pmod{v_0}$$

i.e. $(\delta-\delta^*)\equiv 0 \pmod{v_0}$. Since δ^* is a unit at v_0, we can divide to conclude that

$$\delta/\delta^* \equiv 1 \pmod{v_0}.$$

But since it's also an ℓth root of unity with $v_0 \nmid \ell$, it's actually equal to 1. □

3.4.2. It's often important to know a statement of the form: for $U \supset V$, $S_{k,\chi}(A)^V$ is free over $A[U/V]$. This tells you that as you varying V, the growth of the space of automorphic forms (and congruences between such) is as big as it could be. The next lemma is a precise form of this statement.

Lemma 3.4.4. *Suppose $U \triangleright V$, and U/V has ℓ-power order and U is sufficiently small for ℓ. If A is a local $\mathcal{O}$-algebra, then $S_{k,\chi}(A)^V$ is a finite free module over*

$$A[U\cdot(\mathbf{A}_F^\infty)^\times/V\cdot(\mathbf{A}_F^\infty)^\times] = A[U/V(U\cap(\mathbf{A}_F^\infty)^\times)],$$

and the map $S_{k,\chi}(A)^V \xrightarrow{\mathrm{Tr}_{U/V}} S_{k,\chi}(A)^U$ factors through an isomorphism

$$S_{k,\chi}(A)^V/U \xrightarrow{\sim} S_{k,\chi}(A)^U.$$

Proof. The second part is automatic from the freeness. For the first, we write

$$S_{k,\chi}^V = \bigoplus_{g\in D^\times\backslash(D\otimes\mathbf{A}_F^\infty)^\times/V\cdot(\mathbf{A}_F^\infty)^\times} \Big(\bigotimes_\tau \mathrm{Sym}^{n_\tau}(A^2)\otimes(\det)^{m_\tau}\Big)^{V\cdot(\mathbf{A}_F^\infty)^\times\cap g^{-1}D^\times g}$$

The key is to split this up according to U:

$$\bigoplus_{g\in D^\times\backslash(D\otimes\mathbf{A}_F^\infty)^\times/U\cdot(\mathbf{A}_F^\infty)^\times}\ \bigoplus_{h\in D^\times\backslash D^\times gU\cdot(\mathbf{A}_F^\infty)^\times/V\cdot(\mathbf{A}_F^\infty)^\times} \Big(\bigotimes_\tau \mathrm{Sym}^{n_\tau}(A^2)\otimes(\det)^{m_\tau}\Big)^{V\cdot(\mathbf{A}_F^\infty)^\times\cap h^{-1}D^\times h} \tag{3.4.1}$$

We have

$$F^\times \subset V\cdot(\mathbf{A}_F^\infty)^\times\cap h^{-1}D^\times h \subset U\cdot(\mathbf{A}_F^\infty)^\times \cap h^{-1}D^\times h$$

but our assumptions imply both that $[U\cdot(\mathbf{A}_F^\infty)^\times\cap h^{-1}D^\times h : F^\times]$ is coprime to ℓ and $[U\cdot(\mathbf{A}_F^\infty)^\times\cap h^{-1}D^\times h : V\cdot(\mathbf{A}_F^\infty)^\times\cap h^{-1}D^\times h]$ is a power of ℓ, so the latter index must be equal to 1.

Next we rewrite the double coset decomposition

$$\begin{aligned} D^\times\backslash D^\times gU\cdot(\mathbf{A}_F^\infty)^\times/V\cdot(\mathbf{A}_F^\infty)^\times &= D^\times\cap g(\mathbf{A}_F^\infty)^\times g^{-1}\backslash gU(\mathbf{A}_F^\infty)^\times/V\cdot(\mathbf{A}_F^\infty)^\times \\ &= g(g^{-1}D^\times g\cap U(\mathbf{A}_F^\infty)^\times\backslash U\cdot(\mathbf{A}_F^\infty)^\times/V\cdot(\mathbf{A}_F^\infty)^\times \\ &= g(g^{-1}D^\times g\cap V(\mathbf{A}_F^\infty)^\times\backslash U\cdot(\mathbf{A}_F^\infty)^\times/V\cdot(\mathbf{A}_F^\infty)^\times \\ &= g(U\cdot(\mathbf{A}_F^\infty)^\times/V\cdot(\mathbf{A}_F^\infty)^\times). \end{aligned}$$

So we can rewrite the sum as

$$\bigoplus_{g \in D^\times \backslash (D \otimes \mathbf{A}_F^\infty)^\times / U \cdot (\mathbf{A}_F^\infty)^\times} \bigoplus_{u \in U \cdot (\mathbf{A}_F^\infty)^\times / V \cdot (\mathbf{A}_F^\infty)^\times} (\bigotimes_\tau \mathrm{Sym}^{n_\tau}(A^2) \otimes (\det)^{m_\tau})^{u^{-1}(U \cdot (\mathbf{A}_F^\infty)^\times \cap g^{-1} D^\times g) u}$$
$$= \bigoplus_{g \in D^\times \backslash (D \otimes \mathbf{A}_F^\infty)^\times / U \cdot (\mathbf{A}_F^\infty)^\times} \bigoplus_{u \in U \cdot (\mathbf{A}_F^\infty)^\times / V \cdot (\mathbf{A}_F^\infty)^\times} u^{-1} (\bigotimes_\tau \mathrm{Sym}^{n_\tau}(A^2) \otimes (\det)^{m_\tau})^{(U \cdot (\mathbf{A}_F^\infty)^\times \cap g^{-1} D^\times g)}$$
$$= \bigoplus_{g \in D^\times \backslash (D \otimes \mathbf{A}_F^\infty)^\times / U \cdot (\mathbf{A}_F^\infty)^\times} A[U \cdot (\mathbf{A}_F^\infty)^\times / V \cdot (\mathbf{A}_F^\infty)^\times] \otimes_A (\bigotimes_\tau \mathrm{Sym}^{n_\tau}(A^2) \otimes (\det)^{m_\tau})^{(U \cdot (\mathbf{A}_F^\infty)^\times \cap g^{-1} D^\times g)}$$

$\square$

Remark 3.4.5. Morally, what's going on is that

$$D^\times \backslash \mathrm{GL}_2(\mathbf{A}_F^\infty) / V(\mathbf{A}_F^\infty)^\times \to D^\times \backslash \mathrm{GL}_2(\mathbf{A}_F^\infty) / U(\mathbf{A}_F^\infty)^\times$$

is an étale cover (with group $U(\mathbf{A}_F^\infty)^\times / V(\mathbf{A}_F^\infty)^\times$) when the things are small enough. The statement is that H^0 upstairs is a free module over H^0 downstairs (since everything is 0-dimensional).

In higher dimensional cases, you have more cohomology to worry about, but you still have a statement at the level of complexes.

3.5. **Hecke operators.** We have an action of $\mathrm{GL}_2(\mathbf{A}_F^{\infty,\ell})$ on $S_{k,\chi}(A)$. When we take $S_{k,\chi}(A)^U$, we lose the group action. But we have something left, which is the action of Hecke operators. So that's what we'll explain next.

Let S be a finite set of primes such that $U \supset \mathrm{GL}_2(\widehat{\mathcal{O}}_F^S)$, where

$$\widehat{\mathcal{O}}_F^S = \prod_{v \notin S} \mathcal{O}_{F,v}.$$

Hence U contains something of the form $\mathrm{GL}_2(\widehat{\mathcal{O}}_F^S) \times (U_S \subset \prod_{v \in S} \mathrm{GL}_2(F_v))$.

3.5.1. If $UgU = \coprod g_i U$ is a double coset, then its action on $f \in S_{k,\chi}(A)^U$ is defined by

$$(UgU) \cdot (\varphi) = \sum g_i \varphi$$

Note that $g_i \varphi$ itself is not fixed by U (it's fixed by $g_i U g_i^{-1}$) but the sum is fixed by U.

Another way of presenting this is:

$$(UgU) \cdot (\varphi) = \frac{1}{dh(U)} \int_{\mathrm{GL}_2(\mathbf{A}_F^\infty)} \mathbf{I}_{UgU}(h) h \cdot \varphi \, dh.$$

Definition 3.5.1. We define the *Hecke algebra* $\mathbf{T}_{k,\chi}^S(U,A) \subset \mathrm{End}_A(S_{k,\chi}(A)^U)$ to be the subalgebra generated by operators T_v associated to the double coset

$$U \begin{pmatrix} \varpi_v & \\ & 1 \end{pmatrix} = \begin{pmatrix} 1 & \\ & \varpi_v \end{pmatrix} U \cup \coprod_{x \in \mathcal{O}_{F,v}/(\varpi_v)} \begin{pmatrix} \varpi_v & \alpha \\ 0 & 1 \end{pmatrix}$$

for $v \notin S$.

3.5.2. *Commutativity.* It's easy to check that the Hecke algebra $\mathbf{T}^S_{k,\chi}(U,A)$ is commutative. This is essentially because the generators T_v are supported at different v, and hence don't interact with each other. It is a finite A-module if A is Noetherian, because it's a subalgebra of the whole endomorphism algebra, which is evidently a finite free A-algebra.

3.5.3. *Change of base.* Recall that for flat B/A, we have $S_{k,\chi}(A)^U \otimes_A B \cong S_{k,\chi}(B)^U$. This induces a map

$$\mathbf{T}^S_{k,\chi}(U,A)\otimes_A B \twoheadrightarrow \mathbf{T}^S_{k,\chi}(U,B).$$

(When you can get a map between Hecke algebras, surjectivity is usually evident, as the generators are defined already in $\mathbf{T}^S_{k,\chi}(U,A)$. It's getting a map at all that can be tricky.) If U is sufficiently small with respect to ℓ, then the same thing is true without the flatness assumption, since we still know that $S_{k,\chi}(A)^U \otimes_A B \cong S_{k,\chi}(B)^U$ in this case (§3.4).

Example 3.5.2. A typical application is: $A=\mathcal{O}$ and $B=\mathcal{O}/\varpi$ the residue field.

Without the hypothesis that U is sufficiently small, $S_{k,\chi}(B)^U$ can be "bigger" than $S_{k,\chi}(A)^U \otimes_A B$, so you have no map of Hecke algebras (relations in the smaller space need not be satisfied in the bigger space).

Lemma 3.5.3. *If $A\twoheadrightarrow B$ and A is noetherian, then the kernel of the map*

$$\mathbf{T}^S_{k,\chi}(U,A)\otimes_A B \twoheadrightarrow \mathbf{T}^S_{k,\chi}(U,B)$$

is nilpotent.

Proof. Let $I=\ker(A\twoheadrightarrow B)$. Suppose $T\in \mathbf{T}^S_{k,\chi}(U,A)$ maps to 0 in $\mathbf{T}^S_{k,\chi}(U,B)$. Then $TS_{k,\chi}(A)\subset IS_{k,\chi}(A)$, hence $T^n\,\mathrm{End}_A(S_{k,\chi}(U,A))\subset I^n\,\mathrm{End}_A(S_{k,\chi}(U,A))$ for all n. By the Artin-Rees Theorem, for $n\gg 0$ we get

$$T^n\in I^n\,\mathrm{End}_A(S_{k,\chi}(U,A))\cap \mathbf{T}^S_{k,\chi}(U,A)\subset I\mathbf{T}^S_{k,\chi}(U,A).$$

This is what it means for T^n to become 0 in $\mathbf{T}^S_{k,\chi}(U,A)\otimes_A B$. $\square$

This implies that the two sides have the same maximal ideals, for example. Now, since $\mathbf{T}^S_{k,\chi}(U,\mathcal{O})$ is a finite $\mathcal{O}$-module, we have

$$\mathbf{T}^S_{k,\chi}(U,\mathcal{O})\cong\bigoplus_{\mathfrak{m}}\mathbf{T}^S_{k,\chi}(U,\mathcal{O})_{\mathfrak{m}}.$$

Then we can also break up

$$S_{k,\chi}(\mathcal{O})^U=\bigoplus_{\mathfrak{m}}S_{k,\chi}(\mathcal{O})^U_{\mathfrak{m}}.$$

3.5.4. We fix $\iota:\overline{\mathbf{Q}}_\ell\cong\mathbf{C}$ so that we can base change to $\mathbf{C}$. Upon doing so, we find that

$$\mathbf{T}^S_{k,\chi}(U,\mathbf{C})\cong\prod_{\mathscr{A}_{k,\chi}(U)}\mathbf{C}\times\prod_{\mathscr{C}_{k,\chi}(U)}\mathbf{C} \tag{3.5.1}$$

where the two products are indexed by:

(1) $\mathscr{A}_{k,\chi}(U)$ runs over cuspidal automorphic representations π of $\mathrm{GL}_2(\mathbf{A}_F)$ satisfying
- $\mathrm{HC}_\tau(\pi_\infty) = \{m_\tau + n_\tau + 1, m_\tau\}$,
- $\chi_\pi = \chi||\cdot||^{-1}$, and
- $(\pi^\infty)^U \neq 0$.

The point is that T_v will act as a scalar on $(\pi_v \otimes |\det|_v^{1/2})^{\mathrm{GL}_2(\mathscr{O}_v)}$, and the isomorphism sends T_v to that scalar.

(2) If $n_\tau = 0$ for all τ, then we get a second product over $\mathscr{C}_{k,\chi}(U)$, which indexes Hecke characters $\phi : \mathbf{A}_F^\times / F^\times \to \mathbf{C}^\times$ satisfying
- $\phi^2 = \chi$ and
- $\phi(\det U) = \{1\}$,

which sends T_v to $(1 + \#k(v))\phi(\varpi_v)$.

Why is the map surjective? Suppose it landed in a subalgebra, which would be defined by two coordinates being equal. That would mean something like: you have two automorphic representations π, π' with the same T_v-eigenvalues for almost all v. But this forces $\pi = \pi'$ by strong multiplicity one.

3.6. **Unramified representations.** In terms of Galois representations, the scalars appearing in (3.5.1) are:

(1) In the first product, T_v is sent to the eigenvalue of T_v on $(\pi_v|\det|_v^{1/2})^{\mathrm{GL}_2(\mathscr{O}_v)}$, which is the trace of $\mathrm{rec}(\pi_v)(\mathrm{Frob}_v)$.

(2) In the second product, T_v is sent to the quantity $(1 + \#k(v))\phi(\varpi_v)$, which coincides with $\mathrm{rec}(\phi_v)(\mathrm{Frob}_v) + (\mathrm{rec}(\phi_v)|\cdot|_v^{-1})(\mathrm{Frob}_v)$.

3.6.1. *Classification of unramified representations.* We want to explain more about this. There is a classification of local representations π_v such that $\pi_v^{\mathrm{GL}_2(\mathscr{O}_v)} \neq 0$, which are called "unramified" or "spherical". They are all of the form $\pi_v \cong \mathrm{Ind}_{B(F_v)}^{\mathrm{GL}_2(F_v)}(\chi_\alpha \times \chi_\beta)$. Here,

- B is the *Borel subgroup*

$$B = \left\{ \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \right\} \subset \mathrm{GL}_2(F_v).$$

- $\chi_\alpha \times \chi_\beta$ is the character

$$\chi_\alpha \times \chi_\beta \begin{pmatrix} a & b \\ & d \end{pmatrix} = \alpha^{v(a)} \beta^{v(d)}.$$

The $\mathrm{GL}_2(F_v)$-representation $\mathrm{Ind}_{B(F_v)}^{\mathrm{GL}_2(F_v)}(\chi_\alpha \times \chi_\beta)$ refers to the *normalized induction* from B to $\mathrm{GL}_2(F_v)$, which is the space of locally constant functions

$$\left\{ \varphi\colon \mathrm{GL}_2(F_v) \to \mathbf{C}\colon \varphi\left(\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} g \right) = (\alpha \# k(v)^{-1/2})^{v(a)} (\beta \# k(v)^{1/2})^{v(d)} \right\}.$$

Remark 3.6.1. If $\chi_{\alpha/\beta} = |\cdot|^{\pm 1/2}$ then this is reducible, and there will be a 1-dimensional irreducible subspace or quotient. If π_v is the local component of a cuspidal representation, then this will never happen.

3.6.2. The Galois representation associated to $\mathrm{Ind}_{B(F_v)}^{\mathrm{GL}_2(F_v)}(\chi_\alpha \times \chi_\beta)$ is unramified, with

$$\mathrm{rec}(\pi_v)(\mathrm{Frob}_v) = \begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}.$$

3.6.3. *Hecke eigenvalues.* Let's think about what the representation $\mathrm{Ind}_{B(F_v)}^{\mathrm{GL}_2(F_v)}(\chi_\alpha \times \chi_\beta)$ looks like. By the Iwasawa decomposition, we have

$$\mathrm{GL}_2(F_v) = B(F_v)\mathrm{GL}_2(\mathcal{O}_v).$$

This makes it clear that there is only a one-dimensional space of functions on $\mathrm{GL}_2(F_v)$ that transform in a prescribed way under the left action of $B(F_v)$ and right action of $\mathrm{GL}_2(\mathcal{O}_v)$. We pin down a generator $\varphi_0 \in \mathrm{Ind}_{B(F_v)}^{\mathrm{GL}_2(F_v)}(\chi_\alpha \times \chi_\beta)$, defined by:

$$\varphi_0\left(\begin{pmatrix} a & b \\ & d \end{pmatrix} k\right) = (\alpha \# k(v)^{-1/2})^{v(a)}(\beta \# k(v)^{1/2})^{v(d)}$$

(in other words, φ_0 is the invariant function normalized to be 1 on $\mathrm{GL}_2(\mathcal{O}_v)$). Let's calculate the Hecke action on this function. We know from the 1-dimensionality of the $\mathrm{GL}_2(\mathcal{O}_v)$-fixed vectors that

$$T_v \varphi = t_v \varphi_0,$$

for some scalar t_v. To calculate it, we can evaluate at 1, as

$$t_v = T_v \varphi_0(1).$$

We explained last time that T_v is associated to the double coset

$$\mathrm{GL}_2(\mathcal{O}_v)\begin{pmatrix} \varpi_v & \\ & 1 \end{pmatrix}\mathrm{GL}_2(\mathcal{O}_v) = \begin{pmatrix} 1 & \\ & \varpi_v \end{pmatrix}\mathrm{GL}_2(\mathcal{O}_v) \cup \coprod_{a \in \mathcal{O}_{F,v}/(\varpi_v)} \begin{pmatrix} \varpi_v & a \\ 0 & 1 \end{pmatrix}\mathrm{GL}_2(\mathcal{O}_v).$$

Hence

$$\begin{aligned} T_v \varphi(1) &= \varphi_0 \begin{pmatrix} 1 & \\ & \varpi_v \end{pmatrix} + \sum_{a \in k(v)} \varphi_0 \begin{pmatrix} \varpi_v & a \\ 0 & 1 \end{pmatrix} \\ &= \beta \# k(v)^{1/2} + \# k(v)(\alpha \# k(v)^{-1/2}) \\ &= \# k(v)^{1/2}(\alpha_v + \beta_v) \\ &= \# k(v)^{1/2}\,\mathrm{Tr}(\mathrm{rec}(\pi_v)(\mathrm{Frob}_v)). \end{aligned}$$

3.7. **Galois representations in the Hecke algebra.** Using $\iota \colon \overline{\mathbf{Q}}_\ell \cong \mathbf{C}$, we can replace all occurrences of $\mathbf{C}$ with $\overline{\mathbf{Q}}_\ell$. Now we have a representation

$$r := \prod r_{\ell,i}(\pi) \times \prod (r_{\ell,i}(\phi) \oplus \epsilon_\ell^{-1} r_{\ell,i}(\phi)) \colon G_F \to \mathrm{GL}_2(\mathbf{T}^S_{k,\chi}(U, \overline{Q}_\ell))$$

which for all $v \in S$, $v \nmid \ell$, has

$$\mathrm{Tr}\, r(\mathrm{Frob}_v) = [\mathrm{Tr}(r_{\ell,\iota}(\pi)(\mathrm{Frob}_v), r_{\ell,\iota}(\phi)(\mathrm{Frob}_v) + \# k(v) r_{\ell,i}(\phi)(\mathrm{Frob}_v)] = T_v$$

in terms of the identification (3.5.1).

Remark 3.7.1. The fact that the traces are valued in the smaller algebra $\mathbf{T}^S_{k,\chi}(U,\mathcal{O})$ will often imply, by general principles, that the Galois representation is even valued in $\mathbf{T}^S_{k,\chi}(U,\mathcal{O})$. Then one can ask questions about congruences, which will introduce more power and subtleties.

3.8. **Generalizations to nebentypus.** The theory discussed above has a version adapted to a "nebentypus" character. Let $\psi\colon U_S \to \mu_\infty(\overline{\mathbf{Q}}_\ell)$ be a character, valued in $\mathcal{O}^*$, such that

$$\psi|_{U_S\cap\mathbf{A}^\times_{F,S}} = \chi|_{U_S\cap\mathbf{A}^\times_{F,S}}.$$

In analogy to Lemma 3.2.5, we define

$$S_{k,\chi}(A)^{U,\psi} = \{\varphi \in S_{k,\chi}(A)\colon u_\ell\varphi(gu) = \psi(u)\varphi(g)\colon u \in U\}.$$

We then get an action of $\mathbf{T}^S_{k,\chi}(U,\psi,A)$ on $S_{k,\chi}(A)^{U,\psi}$. Analogously to (3.5.1), there is a decomposition

$$\mathbf{T}^S_{k,\chi}(U,\psi,\mathbf{C}) = \prod_\pi \times \ldots \tag{3.8.1}$$

3.8.1. Using the theory of congruences, we can relate this Hecke algebra to the one without Nebentypus. We'll explain this now. Suppose U is sufficiently small for ℓ and ψ is of ℓ-power order. We know from Lemma 3.5.3 that there is a map

$$\mathbf{T}^S_{k,\chi}(U,\mathcal{O}_{\overline{\mathbf{Q}}_\ell})\otimes\overline{\mathbf{F}}_\ell \twoheadrightarrow \mathbf{T}^S_{k,\chi}(U,\overline{\mathbf{F}}_\ell)$$

with nilpotent kernel. The analogous fact holds for ψ:

$$\mathbf{T}^S_{k,\chi}(U,\psi,\mathcal{O}_{\overline{\mathbf{Q}}_\ell})\otimes\overline{\mathbf{F}}_\ell \twoheadrightarrow \mathbf{T}^S_{k,\chi}(U,\psi,\overline{\mathbf{F}}_\ell)$$

is a surjection with nilpotent kernel. Since ℓth roots of 1 are just 1 in characteristic ℓ, the targets are the same.

$$\begin{array}{ccc}
\mathbf{T}^S_{k,\chi}(U,\mathcal{O}_{\overline{\mathbf{Q}}_\ell})\otimes\overline{\mathbf{F}}_\ell & & \mathbf{T}^S_{k,\chi}(U,\psi,\mathcal{O}_{\overline{\mathbf{Q}}_\ell})\otimes\overline{\mathbf{F}}_\ell \\
\big\downarrow & & \big\downarrow \\
\mathbf{T}^S_{k,\chi}(U,\overline{\mathbf{F}}_\ell) & = & \mathbf{T}^S_{k,\chi}(U,\psi,\overline{\mathbf{F}}_\ell)
\end{array}$$

3.8.2. Suppose you have a map

$$\theta\colon \mathbf{T}^S_{k,\chi}(U,\overline{\mathbf{Q}}_\ell) \to \overline{\mathbf{Q}}_\ell$$

sending $\mathbf{T}^S_{k,\chi}(U,\mathcal{O}_{\overline{\mathbf{Q}}_\ell}) \to \mathcal{O}_{\overline{\mathbf{Q}}_\ell}$. Then there exists

$$\theta'\colon \mathbf{T}^S_{k,\chi}(U,\psi,\mathcal{O}_{\overline{\mathbf{Q}}_\ell}) \to \overline{\mathbf{Q}}_\ell$$

which is congruent mod $\mathfrak{m}_{\overline{\mathbf{Q}}_\ell}$. Indeed, by reducing θ mod $\mathfrak{m}_{\overline{\mathbf{Q}}_\ell}$ we get a character $\overline{\theta}\colon \mathbf{T}^S_{k,\chi}(U,\mathcal{O}_{\overline{\mathbf{Q}}_\ell})\otimes\overline{\mathbf{F}}_\ell\to\overline{\mathbf{F}}_\ell$. Such a character descends uniquely over any nilpotent kernel, giving

$$\begin{tikzcd}
 & \mathbf{T}^S_{k,\chi}(U,\mathcal{O}_{\overline{\mathbf{Q}}_\ell})\otimes\overline{\mathbf{F}}_\ell \arrow[dl, "\overline{\theta}"'] \arrow[d, two heads] & \mathbf{T}^S_{k,\chi}(U,\psi,\mathcal{O}_{\overline{\mathbf{Q}}_\ell})\otimes\overline{\mathbf{F}}_\ell \arrow[d, two heads] \\
\overline{\mathbf{F}}_\ell & \mathbf{T}^S_{k,\chi}(U,\overline{\mathbf{F}}_\ell) \arrow[l, dashed] \arrow[r, equal] & \mathbf{T}^S_{k,\chi}(U,\psi,\overline{\mathbf{F}}_\ell)
\end{tikzcd}$$

Chasing through the diagram, we get a character

$$\overline{\theta}'\colon \mathbf{T}^S_{k,\chi}(U,\psi,\mathcal{O}_{\overline{\mathbf{Q}}_\ell})\otimes\overline{\mathbf{F}}_\ell\to\overline{\mathbf{F}}_\ell.$$

3.8.3. Then we get an integral lift by the going-down theorem, applied to the finite flat map

$$\begin{tikzcd}
\operatorname{Spec}\mathbf{T}^S_{k,\chi}(U,\psi,\mathcal{O}_{\overline{\mathbf{Q}}_\ell}) \arrow[d] \\
\operatorname{Spec}\mathcal{O}_{\overline{\mathbf{Q}}_\ell}.
\end{tikzcd}$$

Namely, viewing $\ker\overline{\theta}'$ as an ideal of $\mathbf{T}^S_{k,\chi}(U,\psi,\mathcal{O}_{\overline{\mathbf{Q}}_\ell})$ lying over $\mathfrak{m}_{\overline{\mathbf{Q}}_\ell}\subset\mathcal{O}_{\overline{\mathbf{Q}}_\ell}$, we have

$$\begin{tikzcd}
\mathfrak{p} \arrow[r, hook, dashed] \arrow[d, dash, dashed] & \ker\overline{\theta}' \arrow[r, hook] \arrow[d, dash] & \mathbf{T}^S_{k,\chi}(U,\psi,\mathcal{O}_{\overline{\mathbf{Q}}_\ell}) \arrow[d, dash] \\
(0) \arrow[r, hook] & \mathfrak{m}_{\overline{\mathbf{Q}}_\ell} \arrow[r, hook] & \mathcal{O}_{\overline{\mathbf{Q}}_\ell}
\end{tikzcd}$$

so we can find $\mathfrak{p}\subset\ker\overline{\theta}'$ lying over $(0)\subset\mathcal{O}_{\overline{\mathbf{Q}}_\ell}$. Since $\mathbf{T}^S_{k,\chi}(U,\psi,\mathcal{O}_{\overline{\mathbf{Q}}_\ell})/\mathfrak{p}$ is finite and an integral domain over $\mathcal{O}_{\overline{\mathbf{Q}}_\ell}$, they must be equal.

To summarize, we have established the following:

Theorem 3.8.1. *Suppose π is a regular algebraic cuspidal automorphic representation of $\mathrm{GL}_2(\mathbf{A}_F)$. Let U be a sufficiently small for ℓ, open compact subgroup of $\mathrm{GL}_2(\mathbf{A}_F^\infty)$ such that $\pi^U\neq 0$ and let S be a sufficiently large finite set of primes such that $\mathrm{GL}_2(\widehat{\mathcal{O}}_F^S)\subset U$.*

Let ψ be a finite ℓ-power order character of U_S such that $\chi_\pi|_{U_S\cap(\mathbf{A}_F^\infty)^\times}=\psi|_{U_S\cap(\mathbf{A}_F^\infty)^\times}$. If $\overline{r_\ell(\pi)}$ (the mod ℓ representation associated to π) is irreducible, then there exists a regular algebraic cuspidal automorphic representation π' of $\mathrm{GL}_2(\mathbf{A}_F)$ such that

- $\mathrm{HC}(\pi_\infty)=\mathrm{HC}(\pi'_\infty)$,
- $\chi_\pi=\chi_{\pi'}$,
- $\overline{r_\ell(\pi)}\cong\overline{r_\ell(\pi')}$,
- $(\pi')^{U,\psi}\neq 0$.

Indeed, by (3.8.1) the character θ' corresponds to either a π' or a ϕ. But the latter case has a reducible residual representation, so it is ruled out (as the residual representation of θ' is automatically irreducible, being isomorphic to $\overline{r_\ell(\pi)}$.

4. Completion of the reduction steps

We now return to the setting of Theorem 2.1.1. Recall the reductions we have made in §2.4.

4.1. **Setup.** We will now complete the proof of reduction §2.4.4.

4.1.1. Let π be our regular cuspidal automorphic representation. Let S be the finite set of places such that π_v is ramified exactly when $v \in S$. In §2.4 we have already made some reductions so that we can assume that $\pi_v^{\mathrm{Iw}_v} \neq \{0\}$ and that $\# k(v) \equiv 1 \pmod{\ell}$ for all $v \in S$. Recall that

$$\mathrm{Iw}_v = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathcal{O}_{F,v}) \mid v(c) > 0 \right\},$$

and the condition that $\pi_v^{\mathrm{Iw}_v} \neq \{0\}$ when π_v is ramified tells us exactly that

$$\mathrm{rec}(\pi_v) \simeq \left(\rho, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right) \quad \text{with } \rho \text{ unramified.}$$

4.1.2. We want to apply Theorem 3.8.1 to obtain a π' with $\overline{r_\ell(\pi)} \simeq \overline{r_\ell(\pi')}$ such that π'_v is *potentially unramified* for all v, i.e. $\mathrm{rec}(\pi'_v) = (\rho, 0)$ for some ρ (possibly ramified). We will choose some auxiliary prime $v_0 \notin S$ with $v_0 \nmid 2\ell$, and let $S' = S \cup \{v_0\}$. Let

$$U = \left(\prod_{v \in S} \mathrm{Iw}_v \right) \times \mathrm{Iw}^1_{v_0} \times \mathrm{GL}_2(\widehat{\mathcal{O}_F^{S'}})$$

where

$$\mathrm{Iw}^1_{v_0} = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid v_0(c) > 0, v_0(\frac{a}{d} - 1) > 0 \right\}.$$

This implies that U is sufficiently small for ℓ.

4.1.3. We choose a character

$$\psi = \prod_{v \in S} \psi_v : \prod_{v \in S} \mathrm{Iw}_v \to \mu_{\ell^\infty},$$

defined by

$$\psi_v \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \right) = \psi'_v(\frac{a}{d}) \tag{4.1.1}$$

with ψ' some non-trivial ℓ-power order character (note that we reduced to $\ell \mid \# k(v)$ for all $v \in S$, so some such ψ' exists).

Now, Theorem 3.8.1 tells us that there exists π', a regular algebraic cuspidal automorphic representation of $\mathrm{GL}_2(\mathbf{A}_F)$ such that:

- π' is unramified away from S'
- $\chi_{\pi'} = \chi_\pi$
- $\mathrm{HC}_\tau(\pi'_\infty) = \mathrm{HC}_\tau(\pi_\infty)$ for all τ

- $\overline{r_\ell(\pi)} \simeq \overline{r_\ell(\pi')}$
- $(\pi')^{U,\psi} \neq \{0\}$

4.2. **Local representations with Iwahori fixed vector.** We have a classification of the possible local representations:

Proposition 4.2.1 (Classification of representations with Iwahori fixed vectors)**.** *If π' is a cuspidal automorphic representation of $\mathrm{GL}_2(\mathbf{A}_F)$*[7] *and $(\pi'_v)^{\mathrm{Iw}^1_v} \neq \{0\}$, then either:*

(i) $\mathrm{rec}(\pi'_v) = (\chi_1 \oplus \chi_2, 0)$ *with χ_i a tamely ramified character of W_{F_v}. Equivalently,*
$$\pi^1_v = \mathrm{Ind}_{B(F_v)}^{\mathrm{GL}_2(F_v)}\big((\chi_1 \circ \mathrm{Art}), (\chi_2 \circ \mathrm{Art})\big).$$
(ii) $\mathrm{rec}(\pi'_v) = (\epsilon_\ell \chi \oplus \chi, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix})$ *with χ a tamely ramified character of W_{F_v} (and ϵ_ℓ the cyclotomic character). Equivalently, π'_v is a subquotient of*
$$\mathrm{Ind}_{B(F_v)}^{\mathrm{GL}_2(F_v)}\big((\chi \circ \mathrm{Art}), (\chi \circ \mathrm{Art})|\cdot|_v)\big).$$

Let $v \in S'$. Since Iw^1_v is the kernel of ψ_v inside Iw_v and $(\pi^1_v)^{\mathrm{Iw}_v,\psi_v} \neq (0)$ for $v \in S'$, we have $(\pi'_v)^{\mathrm{Iw}^1_v} \neq \{0\}$ for all $v \in S$. Hence Proposition 4.2.1 applies. To prove that part (i) of the reduction in §2.4.4 can be arranged, we need to rule out possibility (ii). So let's analyze that case.

4.2.1. We have:
$$\begin{aligned}&\mathrm{Ind}_{B(F_v)}^{\mathrm{GL}_2(F_v)}\big((\chi \circ \mathrm{Art}), (\chi \circ \mathrm{Art})|\cdot|_v)\big)\\ &= \big\{\varphi \colon \mathrm{GL}_2(F_v) \to \mathbf{C} \mid \varphi(bgk) = (\chi \times \chi)(b)\psi_v(k)\varphi(g) \text{ for all } b \in B(F_v), k \in \mathrm{Iw}_v\big\}.\end{aligned}$$
We can apply the *Cartan decomposition* for $\mathrm{GL}_2(F_v)$, which says that $\mathrm{GL}_2(F_v) = B(F_v)\mathrm{GL}_2(\mathscr{O}_{F_v})$. This lets us rewrite the above as
$$\begin{aligned}&\mathrm{Ind}_{B(F_v)}^{\mathrm{GL}_2(F_v)}\big((\chi \circ \mathrm{Art}), (\chi \circ \mathrm{Art})|\cdot|_v)\big)\\ &= \big\{\varphi \colon \mathrm{GL}_2(k(v)) \to \mathbf{C} \mid \varphi(bgk) = (\chi \times \chi)(b)\psi_v(k)\varphi(g) \text{ for all } b \in B(k(v)), k \in \mathrm{Iw}_v\big\}.\end{aligned}$$

4.2.2. Combining this with the *Bruhat decomposition*
$$\mathrm{GL}_2(k(v)) = B(k_v) \coprod B(k(v)) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} B(k(v))$$
exhibits $\mathrm{Ind}_{B(F_v)}^{\mathrm{GL}_2(F_v)}\big((\chi \circ \mathrm{Art}), (\chi \circ \mathrm{Art})|\cdot|_v)\big)$ as a direct sum of two pieces:

(1) The space of functions $\varphi \colon B(k(v)) \to \mathbf{C}$ such that
$$\varphi(bgk) = (\chi \times \chi)(b)\psi_v(k)\varphi(g).$$
In particular, a necessary condition for this to be non-zero is that
$$\chi(b) := (\chi \times \chi)(b) = \psi_v(b) \text{ for all } b \in B(k(v)).$$

[7] We only need to assume that the local representation π_v has certain nice properties that are always satisfied by local parts of cuspidal automorphic representations - e.g. being smooth and admissible

Applying this to

$$b = \begin{pmatrix} x & \\ & y \end{pmatrix} \Longrightarrow \chi(xy) = \psi'_v(x/y)$$

by the choice of the auxiliary character ψ in (4.1.1), which forces ψ to be trivial – contradiction.

(2) The space of functions

$$\varphi \colon B(k(v)) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} B(k(v)) \to \mathbf{C}$$

such that

$$\varphi(bgk) = (\chi \times \chi)(b)\psi_v(k)\varphi(g).$$

In this case, we have

$$\varphi\left(\begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right) = \chi(xy)\varphi\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right) = \psi'_v(y/x)\varphi\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right)$$

which again leads to a contradiction.

4.3. **Summary of progress.** Now, we have reduced Theorem 2.1.1 to the following:

Theorem 4.3.1. *Fix some* $\iota \colon \overline{\mathbf{Q}_\ell} \xrightarrow{\sim} \mathbf{C}$. *Suppose* $\mathbf{F}$ *is totally real and* $2 \mid [F : \mathbf{Q}]$. *Suppose* $r \colon G_F \to \mathrm{GL}_2(\mathscr{O})$ *is a regular algebraic* ℓ*-adic representation, where* $\mathscr{O} = \mathscr{O}_L$ *for a finite extension* $L/\mathbf{Q}_\ell$, *and* $\tau F \subseteq L$ *for all* $\tau \colon F \hookrightarrow \overline{\mathbf{Q}_\ell}$, *and* $\mathscr{O}/\lambda = \mathbf{F}$. *Suppose that there exists a regular algebraic cuspidal automorphic representation* π *of* $\mathrm{GL}_2(\mathbf{A}_F)$ *such that:*

- $\overline{r} := r \mod \lambda \simeq \overline{r_\ell(\pi)}$,
- $\mathrm{HT}_\tau(r_\ell(\pi)) = \mathrm{HT}_\tau(r)$ *for all* $\tau \colon F \hookrightarrow \overline{\mathbf{Q}_\ell}$,
- $r_\ell(\chi_\pi) = \det r$.

Furthermore, suppose that:

(1) $\overline{r}|_{G_{F(\zeta_\ell)}}$ *is irreducible.*
(2) For all v, π_v *and* $\mathrm{WD}(r|_{G_{F_v}})^{\mathrm{ss}}$ *are unramified, and that* $\#k(v) \equiv 1 \pmod{\ell}$ *for any* v *such that* $\mathrm{WD}(r|_{G_{F_v}})$ *is ramified.*
(3) ℓ *is unramified in* F *and* $\mathrm{HT}_\tau(r) \subseteq \{0, 1, 2, \dots, \ell - 2\}$ *for all* τ.

Then r *is automorphic.*

We will first prove Theorem 4.3.1 under the simplifying assumption that $\mathrm{WD}(r|_{G_{F_v}})$ is itself unramified for all v. This is sometimes called the "minimal case". In practice, the assumption on ramification is frequently too restrictive to be useful, but it demonstrates the main ideas of the proof in general.

5. Setup for automorphy lifting

5.1. **Taylor-Wiles primes.** Now, if $\widetilde{\mathrm{Frob}}_{v_0} \in W_{F_{v_0}}$ lifts Frob_{v_0}, then $r_\ell(\pi')(\widetilde{\mathrm{Frob}}_{v_0})$ has eigenvalues α and $\alpha \cdot (\#k(v_0))$ for some α. Thus, $\overline{r_\ell(\pi)}(\mathrm{Frob}_{v_0})$ has eigenvalues α and $\#k(v_0)\alpha$.

Hence

$$\begin{aligned}(\operatorname{tr}\overline{r_\ell(\pi)}(\mathrm{Frob}_{v_0}))^2 &= \alpha^2(1+\#k(v_0))^2\\ &= [\det\overline{r_\ell(\pi)}(\mathrm{Frob}_{v_0})](1+\#k(v_0))(1+\#k(v_0)^{-1})\\ &= [\det\overline{r_\ell(\pi)}(\mathrm{Frob}_{v_0})](1+\overline{\epsilon_\ell(\mathrm{Frob}_{v_0})})(1+\overline{\epsilon_\ell^{-1}(\mathrm{Frob}_{v_0})}).\end{aligned}$$

If there is some $\sigma\in G_F$ such that:

$$(\operatorname{tr}\overline{r_\ell(\pi)}(\sigma))^2 \not\equiv [\det\overline{r_\ell(\pi)}(\sigma)](1+\overline{\epsilon_\ell(\sigma)})(1+\overline{\epsilon_\ell^{-1}(\sigma)}) \tag{5.1.1}$$

then by the Chebotarev density theorem, then there will be infinitely many $v\notin S\cup\{\ell\}$ such that this congruence doesn't hold for $\sigma=\mathrm{Frob}_v\in G_{F_v}$.

5.1.1. We want to rule out the possibility that the congruence

$$(\operatorname{tr}\overline{r_\ell(\pi)})^2 \equiv [\det\overline{r_\ell(\pi)}](1+\overline{\epsilon_\ell})(1+\overline{\epsilon_\ell^{-1}}) \tag{5.1.2}$$

holds on all of G_F. We will need the following group-theoretic lemma:

Lemma 5.1.1. *If $G\subseteq \mathrm{GL}_2(\overline{\mathbf{F}}_\ell)$ is a subgroup such that for all $g\in G$, the characteristic polynomial of g is $(x-\alpha)^2$ for some α, then up to conjugation, we may take $G\subseteq \overline{\mathbf{F}}_\ell^\times\begin{pmatrix}1 & \overline{\mathbf{F}}_\ell\\ 0 & 1\end{pmatrix}$ and G is abelian.*

On $G_{F(\zeta_\ell)}$, (5.1.2) implies

$$(\operatorname{tr}\overline{r_\ell(\pi)})^2 = 4\det\overline{r_\ell(\pi)},$$

so $\overline{r_\ell}(\sigma)$ has eigenvalues α,β with $(\alpha+\beta)^2=4\alpha\beta$, so $(\alpha-\beta)^2=0$ so $\alpha=\beta$. Hence Lemma 5.1.1 implies in this case that

$$\overline{r_\ell(\pi)}(G_{F(\zeta_\ell)}) \subseteq \overline{\mathbf{F}}_\ell^\times\begin{pmatrix}1 & \overline{\mathbf{F}}_\ell\\ 0 & 1\end{pmatrix}.$$

This leads to two possibilities:

- $\overline{r_\ell(\pi)}(G_{F(\zeta_\ell)})\subseteq \overline{\mathbf{F}}_\ell^\times$, and therefore $\overline{r_\ell(\pi)}(G_F)$ is abelian, so $\overline{r_\ell(\pi)}$ is reducible, giving a contradiction.
- $\overline{r_\ell}(\pi)(G_{F(\zeta_\ell)})$ fixes a unique line $L\subseteq \overline{\mathbf{F}}_\ell^2$. As $G_{F(\zeta_\ell)}$ is normal in G_L, we see that G_F also fixes this line. But this implies that $\overline{r_\ell(\pi)}$ is reducible, again giving us a contradiction.

5.1.2. We make the following definition.

Definition 5.1.2. Suppose $\overline{r}: G_F\to \mathrm{GL}_2(\overline{\mathbf{F}}_\ell)$ is a representation.

(1) We call v *nearly harmless* for $\overline{r}$ if
- $v\nmid \ell$,
- $\overline{r}$ is unramified at v, and
- if α,β are the roots of the characteristic polynomial of $\overline{r}(\mathrm{Frob}_v)$ then $\alpha/\beta\neq \#k(v)^{\pm 1}$, i.e. (5.1.1) holds for $\sigma=\mathrm{Frob}_v$.

(2) We call v *harmless* if it is nearly harmless and $\#k(v)\not\equiv 1 \pmod{\ell}$.

As discussed above, these conditions are frequently satisfied:

Lemma 5.1.3. *Assume $\ell \neq 2$. Then:*

(1) If $\overline{r}$ is irreducible, there exists infinitely many nearly harmless primes for $\overline{r}$.
(2) If either $[F(\zeta_\ell):F] > 2$ and $\overline{r}$ is irreducible or $\overline{r}|_{G_{F(\zeta_\ell)}}$ is irreducible, then there exist infinitely many harmless primes for $\overline{r}$.

The proof uses the following classification of the possible projective images of $\overline{r}$:

Lemma 5.1.4. *If $\overline{r}$ is irreducible, the projective image of $\overline{r}$ via the map $\mathrm{GL}_2(\overline{\mathbf{F}_\ell}) \twoheadrightarrow \mathrm{PGL}_2(\overline{\mathbf{F}_\ell})$ is, up to conjugation, one of the following possibilities:*

- A_4,
- S_4,
- A_5,
- $\mathrm{PSL}_2(\mathbf{F}_{\ell^r})$,
- $\mathrm{PGL}_2(\mathbf{F}_{\ell^r})$,
- *A subgroup of the "normalizer of the split Cartan,"* $N_{\mathrm{PGL}_2}(T)$.

5.1.3. Let L be a finite extension of $\overline{\mathbf{Q}_\ell}$, $\mathscr{O} = \mathscr{O}_L$ its ring of integers, and $\mathbf{F} = \mathscr{O}/\lambda$ its residue field.

Lemma 5.1.5. *Suppose A is a ($\mathfrak{m}_A$-adically) complete noetherian local $\mathscr{O}$-algebra with residue field $\mathbf{F}$ and that $r: G_F \to \mathrm{GL}_2(A)$ is a continuous representation. If v is harmless for $\overline{r} := r \pmod{\mathfrak{m}}_A$, then r is unramified at v.*

Here, by saying that A is a local $\mathscr{O}$-algebra, we mean that $\mathscr{O} \to A$ is a local homomorphism: $\mathfrak{m}_A$ lies over λ. Furthermore, we require $\mathscr{O}_\lambda \to A/\mathfrak{m}_A$ to induce an isomorphism of residue fields $\mathbf{F} \xrightarrow{\sim} A/\mathfrak{m}_A$. Examples of such A include $\mathscr{O}$, and $\mathscr{O}[[x]]$.

Proof. Let $I \subseteq \mathfrak{m}_A$ be the ideal of A generated by the entries of $r(\sigma)$–Id as σ runs over the inertia subgroup I_{F_v}. We want to prove that $I = \{0\}$ by showing $\mathfrak{m}_A I = I$, which suffices by Nakayama's lemma. We can replace A by $A/\mathfrak{m}_A I$, so without loss of generality we may assume that $\mathfrak{m}_A I = \{0\}$.

There exists some unramified $\widetilde{r}: G_{F_v}/I_{F_v} \to \mathrm{GL}_2(A)$ such that $\widetilde{r} \pmod{I} = r \pmod{I}$, defined by letting $\widetilde{r}(\mathrm{Frob}_v)$ be any lift of $(r \mod I)(\mathrm{Frob}\ v)$. Writing down the condition that $r(\sigma\tau) = r(\sigma)r(\tau)$, we find that for any $\sigma \in G_{F_v}$, we have

$$r(\sigma) = (1+\alpha(\sigma))\cdot\widetilde{r}(\sigma) = (1+\alpha(\sigma)+\mathrm{ad}(\widetilde{r}(\sigma))(\alpha(\tau)))\cdot\widetilde{r}(\sigma\tau).$$

This shows that α defines a cohomology class in $H^1(G_{F_v}, M_{2\times 2}(I))$, where G_{F_v}/I_{F_v} acts on $M_{2\times 2}(I)$ through $\widetilde{r}$ and the adjoint representation.

Now, we have an exact sequence of cohomology:

$$\begin{aligned} 0 \to H^1(G_{F_v}/I_{F_v}, M_{2\times 2}(I)) \to H^1(G_{F_v}, M_{2\times 2}(I)) \to & H^1(I_{F_v}, M_{2\times 2}(I))^{G_{F_v}} \\ &= \mathrm{Hom}_{G_{F_v}}(I_{F_v}, M_{2\times 2}(I)) \\ &= \mathrm{Hom}(\mathbf{Z}_\ell(1), M_{2\times 2}(I)). \end{aligned}$$

(To see the last equality, note that $M_{2\times 2}(I)$ is an ℓ-group, so a map from I_{F_v} to it factors through the maximal pro-ℓ quotient. For $v \mid p$, we have a sequence of normal subgroups $G_{F_v} \triangleright I_{F_v} \triangleright P_{F_v}$ with P_{F_v} the *wild inertia group*. It is a pro-p group. We have $G_{F_v}/I_{F_v} \simeq G_{k(v)} = \langle \mathrm{Frob}_v \rangle \simeq \widehat{\mathbf{Z}}$, and $I_{F_v}/P_{F_v} = \prod_{\ell \neq p} \mathbf{Z}_\ell(1)$. The meaning of writing $\mathbf{Z}_\ell(1)$

instead of $\mathbf{Z}_\ell$ here is that if $\sigma \in G_{F_v}$, its action by conjugation on I_{F_v} corresponds to the map $\tau \mapsto \epsilon_\ell(\sigma)\tau$ on $\mathbf{Z}_\ell$.)

Now, we have $\mathrm{Hom}_{G_{F_v}}(\mathbf{Z}_\ell(1), M_{2\times 2}(I)) = M_{2\times 2}(I)(-1)^{G_{F_v}}$. We will see that there are no invariants. Say $\overline{r}(\mathrm{Frob}_v)$ has eigenvalues α, β. Then $\mathrm{ad}\,\overline{r}$ has eigenvalues 1 (with multiplicity 2) and $(\alpha/\beta)^{\pm 1}$, so the assumption that v is harmless implies that

$$H^1(I_{F_v}, M_{2\times 2}(I))^{G_{F_v}} = (\mathrm{ad}\overline{r})(-1)^{G_{F_v}} \otimes_{\mathbf{F}} I = \{0\}.$$

The long exact sequence in cohomology them implies that there exists a cocycle $\beta \in Z^1(G_{F_v}/I_{F_v}, M_{2\times 2}(I))$ and $a \in M_{2\times 2}(I)$ such that

$$\alpha(\sigma) - \mathrm{ad}(\sigma)(a) + a = \beta(\sigma) \text{ for all } \sigma \in G_{F_v}.$$

Now we have:

$$\begin{aligned}(1+a)r(\sigma)(1+a)^{-1} &= (1+a)(1+\alpha(\sigma))\widetilde{r}(\sigma)(1-a)\\ &= (1+a+\alpha(\sigma) - \mathrm{ad}\widetilde{r}(\sigma)(a))\widetilde{r}(\sigma)\\ &= (1+\beta(\sigma))\widetilde{r}(\sigma)\end{aligned}$$

This is trivial on I_{F_v}, so r is trivial on I_{F_v} as well. □

5.2. **The Galois deformation ring.** We will consider a functor parametrizing certain deformations of the Galois representation $\overline{r}$, which turns out to be representable. More specifically, consider the category $\mathscr{C}_{\mathscr{O},\mathbf{F}}$ consisting of complete noetherian local $\mathscr{O}$-algebras with residue field $\mathbf{F}$.

Definition 5.2.1. The *Galois deformation functor* from $\mathscr{C}_{\mathscr{O},\mathbf{F}}$ to the category of sets is defined by sending an $\mathscr{O}$-algebra A to the set of equivalence classes of continuous representations $\rho\colon G_F \to \mathrm{GL}_2(A)$ such that:

(1) $\overline{\rho} = \rho \pmod{\mathfrak{m}}_A = \overline{r}$.
(2) ρ is unramified at all $v \nmid \ell$.
(3) For all $v \mid \ell$, $\rho|_{G_{F_v}}$ is *Fontaine-Laffaille* ("FL").
(4) $\det\rho = \det r$.

The equivalence relation on such ρ is defined by setting $\rho \sim \rho'$ if $\rho = a\rho' a^{-1}$ for some $a \in \mathrm{GL}_2(A)$ with $a \equiv \mathrm{Id} \pmod{\mathfrak{m}_A}$.

Proposition 5.2.2. *The Galois deformation functor is representable by a ring $R^{\mathrm{univ}} \in \mathscr{C}_{\mathscr{O},\mathbf{F}}$ and a $r^{\mathrm{univ}}\colon G_F \to \mathrm{GL}_2(R^{\mathrm{univ}})$ satisfying conditions (1)-(4) in Definition 5.2.1. The universal property of $(R^{\mathrm{univ}}, \rho^{\mathrm{univ}})$ is that if $\rho\colon G_F \to \mathrm{GL}_2(A)$ satisfies the conditions (1)-(4), then there exists a unique local $\mathscr{O}$-algebra homomorphism $R^{\mathrm{univ}} \to A$ which sends ρ^{univ} to a representation equivalent to ρ.*

The proof of this proposition is essentially just algebra. The key inputs are the following:

- $Z_{\mathrm{GL}_2(\mathbf{F})}(\mathrm{Im}(\overline{r})) = \mathbf{F}^\times$
- The FL objects are closed under the formation of subobjects and quotient objects.

- For all number fields F' and all finite sets of prime S of F', if F'_S/F' is the maximal pro-ℓ extension of F' unramified outside of S, then $\mathrm{Gal}(F'_S/F_S)$ is topologically finitely generated.

The last condition shows that R^{univ} is noetherian.

5.3. **Fontaine-Laffaille theory.** We still need to define what it means for $\rho|_{G_{F_v}}$ to be *Fontaine-Lafaille*. We start by defining a category of "semi-linear" modules, which Fontaine and Lafaille related to a category of ℓ-adic representations of G_{F_v} when $v \mid \ell$.

Definition 5.3.1. Let $v \mid \ell$ with F_v unramified over $\mathbf{Q}_\ell$. We define a category $\mathscr{MF}_v$, whose objects are finite modules M over $\mathscr{O}_{F_v} \otimes_{\mathbf{Z}_\ell} \mathscr{O}$ together with:

(1) A decreasing filtration by $\mathscr{O}_{F_v} \otimes_{\mathbf{Z}_\ell} \mathscr{O}$-submodules $\mathrm{Fil}^i(M)$ such that $\mathrm{Fil}^0(M) = M$ and $\mathrm{Fil}^{\ell-1}(M) = 0$.
(2) Maps $\Phi^i : \mathrm{Fil}^i(M) \to M$ which are $\mathrm{Frob}_\ell^{-1} \otimes 1$-linear and such that $\Phi^i|_{\mathrm{Fil}^{i+1}M} = \ell\Phi^{i+1}$.[8]
(3) $\sum_i \mathrm{Im}(\Phi^i) = M$.

The maps in the category $\mathscr{MF}_v$ are $\mathscr{O}_{F_v} \otimes_{\mathbf{Z}_\ell} \mathscr{O}$-linear maps which respect the filtration and commute with the Φ^i's.

Proposition 5.3.2. *$\mathscr{MF}_v$ is an abelian category.*

This is remarkable because filtered modules typically do not form an abelian category. The relationship with Galois representations comes from:

Proposition 5.3.3. *There exists a covariant functor $\mathbf{G}_{F_v}$ from $\mathscr{MF}_v$ to the category of finite $\mathscr{O}$-modules with a continuous action of G_{F_v} such that:*

- *$\mathbf{G}_{F_v}$ is fully faithful and exact.*
- *The essential image of $\mathbf{G}_{F_v}$ is closed under direct sums, sub-objects, and quotients.*
- *For $M \in \mathscr{MF}_v$,*

$$\mathrm{length}_{\mathscr{O}}(M) = [F_v : \mathbf{Q}_\ell]\mathrm{length}_{\mathscr{O}}(\mathbf{G}_{F_v}(M)).$$

- *If Λ is a torsion-free finitely generated $\mathscr{O}$-module with a continuous action of G_{F_v}, then Λ is in the image of $\mathbf{G}_{F_v}$ if and only if $\Lambda \otimes_{\mathscr{O}} L$ is* crystalline *(equivalently, it is de Rham and* $\mathrm{WD}(\Lambda \otimes_{\mathscr{O}} L)$ *is unramified) and*

$$\mathrm{HT}_\tau(\Lambda \otimes_{\mathscr{O}} L) \subseteq \{0, \dots, \ell - 2\} \text{ for all } \tau.$$

The proofs of both of the preceding propositions can be found in [FL92].

Remark 5.3.4. The Galois representations which are (conjecturally) associated to motives and algebraic automorphic representations are always de Rham, and this restriction is needed to make many theorems work. The condition of being de Rham is a characteristic 0 condition, obtained by considering Galois actions on vector spaces over $\overline{\mathbf{Q}}_\ell$. But when one wants to make Galois deformation arguments, it is necessary to consider torsion coefficients. In general, it is difficult and highly non-explicit to impose the de Rham condition for Galois representations with torsion coefficients - the Fontaine-Laffaille case is one of the few where a direct construction is feasible.

[8] Here, Frob_ℓ^{-1} is the arithmetic Frobenius generating $\mathrm{Gal}(F_v/\mathbf{Q}_\ell)$, defined by requiring $\mathrm{Frob}_\ell^{-1} \alpha \cong \alpha^\ell \pmod{v}$ for any $\alpha \in \mathscr{O}_{F_v}$. The $\mathrm{Frob}_\ell^{-1} \otimes 1$-linearity says exactly that $\Phi^i((\alpha \otimes \beta)m) = (\mathrm{Frob}_\ell^{-1} \alpha \otimes \beta)\Phi^i(m)$.

5.4. **The map $R \to \mathbf{T}$.** Now, choose a prime v_0 which is harmless for $\overline{r}$. Consider the map sending $A \in \mathscr{C}_{\mathscr{O},\mathbf{F}}$ (notation as in Definition 5.2.1) to the set of $\rho\colon G_F \to \mathrm{GL}_2(A)$ such that:

- ρ is unramified outside of ℓv_0
- $\overline{\rho} = \overline{r}$.
- ρ is FL above ℓ.
- $\det\rho = \det r$

This functor is represented by a ring $R^{\mathrm{univ}}_{\{v_0\}}$. By Lemma 5.1.5, the above conditions imply that ρ is unramified at v_0, so we actually get an isomorphism $R^{\mathrm{univ}}_{\{v_0\}} \xrightarrow{\sim} R^{\mathrm{univ}}$ which is compatible with the universal deformations.

Choose a finite set of primes S containing v_0 and all primes above ℓ.

5.4.1. *Spaces of automorphic forms.* Write the Hodge-Tate weights of r as

$$\mathrm{HT}_\tau(r) = \{1 + n_\tau + m_\tau, m_\tau\} \quad \text{with} \quad n_\tau \geq 0 \text{ and } n_\tau, m_\tau \in \mathbf{Z}.$$

Then we have $\mathrm{HT}_\tau(\det r) = \{1 + n_\tau + 2m_\tau\}$. Since F is totally real, $1 + n_\tau + 2m_\tau =: w$ is independent of τ. We will look at automorphic forms (in the sense of §3.3) $S_{k,\chi}(U, \mathscr{O})$ with

- weight $k = \{(n_\tau, m_\tau)\}$,
- algebraic Großencharacter χ satisfying $r_\ell(\chi) = (\det r)\epsilon_\ell$, and
- open compact subgroup

  $$U = \mathrm{GL}_2(\widehat{\mathscr{O}_F}^{v_0}) \times \mathrm{Iw}^1_{v_0} \subseteq \mathrm{GL}_2(\mathbf{A}_K^\infty)$$

  Recall that $\mathrm{Iw}^1_{v_0} = \left\{ g \in \mathrm{GL}_2(\mathscr{O}_{F,v_0}) \mid g \equiv \begin{pmatrix} x & y \\ 0 & x \end{pmatrix} \pmod{v_0} \right\}$. This ensures that U is sufficiently small for v_0.

5.4.2. *The Hecke algebra.* Look at the decomposition of the Hecke algebra for $S_{k,\chi}(U, \mathscr{O})$ given by (3.5.1):

$$\mathbf{T}^S_{k,\chi}(U, \mathscr{O}) \otimes \overline{\mathbf{Q}_\ell} \xrightarrow{\sim} \prod_{\pi \in \mathscr{A}_{k,\chi}(U)} \overline{\mathbf{Q}_\ell} \times \prod_{\phi \in \mathscr{C}_{k,\chi}(U)} \overline{\mathbf{Q}_\ell}.$$

This restricts to a map

$$\mathbf{T}^S_{k,\chi}(U, \mathscr{O}) \to \prod_{\pi' \in A_{k,\chi}(U)} \mathscr{O}_{L'} \times \prod_{\phi \in C_{k,\chi}(U)} \mathscr{O}_{L'}.$$

Here, L'/L is a sufficiently large finite Galois extension such that the projection of $\mathbf{T}^S_{k,\chi}(U, \mathscr{O})$ to each component of $\mathbf{T}^S_{k,\chi}(U, \mathscr{O}) \otimes \overline{\mathbf{Q}_\ell}$ is contained in $\mathscr{O}_{L'} \subseteq \overline{\mathbf{Q}_\ell}$. In other words, we choose L' such that for each $\pi' \in A_{k,\chi}(U)$, $\phi \in C_{k,\chi}(U)$, the image of $r_\ell(\pi')\colon G_F \to \mathrm{GL}_2(\overline{\mathbf{Q}_\ell})$ is contained in $\mathrm{GL}_2(\mathscr{O}_{L'})$ and the image of $r_\ell(\phi)\colon G_F \to \overline{\mathbf{Q}_\ell}^\times$ is contained in $\mathscr{O}_{L'}^\times$.

5.4.3. *Localizing the Hecke algebra.* For each $\pi \in \mathscr{A}_{k,\chi}(U)$, consider the corresponding map

$$\mathbf{T}^S_{k,\chi}(U,\mathscr{O}) \to \mathscr{O} \twoheadrightarrow \mathbf{F}$$

sending T_v to $\operatorname{tr} \overline{r_\ell(\pi)}(\mathrm{Frob}_v)$. Its kernel is a maximal ideal $\mathfrak{m}$ (it is a proper ideal since π exists). We have that $\mathfrak{m} = \langle \lambda, T_v - \operatorname{tr} r(\mathrm{Frob}_v)\rangle_{v\notin S}$.

Now, we localize $\mathbf{T}^S_{k,\chi}(U,\mathscr{O})$ at $\mathfrak{m}$. We claim that localization at $\mathfrak{m}$ kills the factor $\mathscr{O}_{L'}$ of $\mathbf{T}^S_{k,\chi}(U,\mathscr{O})$ corresponding to some $\phi \in C_{k,\chi}(U)$ whenever $\overline{r} \not\simeq \overline{r_\ell(\phi)} \oplus \overline{r_\ell(\phi)}\overline{\epsilon_\ell}^{-1}$, and that it kills the factor $\mathscr{O}_{L'}$ corresponding to $\pi' \in A_{k,\chi}(U)$ whenever $\overline{r_\ell(\pi')} \not\simeq \overline{r_\ell(\pi)}$.

Our assumption that $\overline{r}$ is irreducible forces $\overline{r} \not\simeq \overline{r_\ell(\phi)} \oplus \overline{r_\ell(\phi)}\overline{\epsilon_\ell}^{-1}$. Thus, there exists some $v \notin S$ such that

$$\operatorname{tr} \overline{r}(\mathrm{Frob}_v) \neq \operatorname{tr}\left(\overline{r_\ell(\phi)} \oplus \overline{r_\ell(\phi)}\overline{\epsilon_\ell}^{-1}\right)(\mathrm{Frob}_v)$$

This follows from the fact that:

$$\prod_{\sigma\in \mathrm{Gal}(L'/L)} \left(T_v -^{\sigma} \operatorname{tr}(r_\ell(\phi) \oplus r_\ell(\phi)\epsilon_\ell^{-1})(\mathrm{Frob}_v)\right) \in \mathbf{T}^S_{k,\chi}(U,\mathscr{O}) - \mathfrak{m}$$

If it were in $\mathfrak{m}$, we would have $\operatorname{tr} r(\mathrm{Frob}_v) T_v -^{\sigma} \operatorname{tr}(r_\ell(\phi) \oplus r_\ell(\phi)\epsilon_\ell^{-1})(\mathrm{Frob}_v)$ would be equal to $\operatorname{tr} r(\mathrm{Frob}_v)$.

What about π' with $\overline{r_\ell(\pi')} \simeq \overline{r_\ell(\pi)}$? If $T \notin \mathfrak{m}$, we would have $\alpha \in \mathscr{O} - \lambda$ such that $T - \alpha \in \mathfrak{m}$. This is a polynomial in the T_v's with $\mathscr{O}$-coefficients. Reducing mod $\mathfrak{m}$ means that we reduce mod λ and replace T_v with $\overline{r}(\mathrm{Frob}_v)$.

In the end, we get a decomposition:

$$\mathbf{T}^S_{k,\chi}(U,\mathscr{O})_{\mathfrak{m}} \hookrightarrow \prod_{\substack{\pi \in A_{k,\chi}(U)\\ \overline{r_\ell(\pi)} \simeq \overline{r}}} \mathscr{O}_{L'}$$

Moreover, we can consider the subring:

$$A := \left\{ (x_\pi) \in \prod_\pi \mathscr{O}_{L'} \mid (x_\pi \mod \lambda') \in \mathbf{F} \text{ is independent of } \pi \right\}$$

We can see that $\mathbf{T}^S_{k,\chi}(U,\mathscr{O})_{\mathfrak{m}}$ actually lands inside A, since the components $\pmod{\lambda}$ at each π are $\overline{r_\ell(\pi)}(\mathrm{Frob}_v) = \overline{r}(\mathrm{Frob}_v)$. Furthermore, A is a complete noetherian local $\mathscr{O}$-algebra with residue field $\mathbf{F}$.

5.4.4. *Galois representations.* Given $r_\ell(\pi)\colon G_F \to \mathrm{GL}_2(\mathscr{O}_{L'})$, we can consider the representation $G_F \to \mathrm{GL}_2(\prod_\pi \mathscr{O}_{L'})$, which factors through $\rho\colon G_F \to \mathrm{GL}_2(A)$. Note: for all $v \notin S$,

- ρ is unramified at v,
- $\operatorname{tr}\rho(\mathrm{Frob}_v)$ is the image of $T_v \in \mathbf{T}^S_{k,\chi}(U,\mathscr{O})_{\mathfrak{m}}$ in A, and
- ρ is FL at all $v \mid \ell$.

Suppose you have a continuous representation of G_{F_v}, for $v \mid \ell$, on a finite $\mathscr{O}$-module M. We defined what it means for M to be Fontaine-Laffaille. Now, if $\rho\colon G_{F_v} \to \mathrm{GL}_2(R)$ for R a complete noetherian local $\mathscr{O}$-algebra with residue field $\mathbf{F}$ (which may very well not be finite over $\mathcal{O}$), we say that ρ is *Fontaine-Laffaille* if for every open ideal $I \subseteq R$, $\overline{r}\colon G_{F_v} \to \mathrm{GL}_2(R/I)$ is Fontaine-Laffaille.

Recall that we had the universal lift $r^{\mathrm{univ}}\colon G_F \to \mathrm{GL}_2(R^{\mathrm{univ}}_{\{v_0\}})$. By the universal property, we have a map $R_{\{v_0\}} \to A$ which takes r^{univ} to ρ. What is the image of this map? We claim that it is exactly $\mathbf{T}^S_{k,\chi}(U,\mathcal{O})_{\mathfrak{m}}$. Indeed, since $R^{\mathrm{univ}}_{\{v_0\}}$ is topologically generated by $\mathrm{tr}\, r^{\mathrm{univ}}(\mathrm{Frob}_v)$ for $v \notin S$, we see that the image is topologically generated by the T_v for $v \notin S$, so it is $\mathbf{T}^S_{k,\chi}(U,\mathcal{O})_{\mathfrak{m}}$.

By looking at the functor parametrizing lifts of $\overline{r}$ which are furthermore unramified at v_0, we get a map $R^{\mathrm{univ}}_{\{v_0\}} \to R^{\mathrm{univ}}$: since v_0 is harmless for ℓ, this is an isomorphism.

What is the kernel of the map $R^{\mathrm{univ}}_{\{v_0\}} \to A$? Equivalently, it is the kernel of the surjection $R^{\mathrm{univ}}_{\{v_0\}} \twoheadrightarrow \mathbf{T}^S_{k,\chi}(U,\mathcal{O})$, and also equal to the intersection:

$$\bigcap_{\substack{\pi \in A_{k,\chi}(U) \\ \overline{r_\ell(\pi)} \simeq \overline{r}}} \ker(R^{\mathrm{univ}} \to \overline{\mathbf{Q}_\ell})$$

Here, the map corresponding to the factor π takes r^{univ} to $r_\ell(\pi)$. Note that this last representation does not depend on S, so we will drop S from the notation. We will see that this kernel is actually $\{0\}$. In other words:

Theorem 5.4.1. *We have an isomorphism* $R^{\mathrm{univ}} \xrightarrow{\sim} R^{\mathrm{univ}}_{\{v_0\}} \xrightarrow{\sim} \mathbf{T}_{k,\chi}(U,\mathcal{O})_{\mathfrak{m}}$.

This theorem suffices to prove Theorem 4.3.1 on automorphy lifting. Indeed, consider the map $\theta\colon R^{\mathrm{univ}} \to \mathcal{O}$ which takes r^{univ} to r. The Theorem tells us that θ can be thought of as a map $\theta'\colon \mathbf{T}_{k,\chi}(U,\mathcal{O})_{\mathfrak{m}} \to \mathcal{O}$ which sends T_v to $\mathrm{tr}\, r(\mathrm{Frob}_v)$. Since $\mathbf{T}_{k,\chi}(U,\mathcal{O})_{\mathfrak{m}}$ sits inside the product $\prod_{\substack{\pi \in A_{k,\chi}(U) \\ \overline{r_\ell(\pi)} \simeq \overline{r}}}$, there is some $\pi \in A_{k,\chi}(U)$ with $\overline{r_\ell(\pi)} \simeq \overline{r}$ such that $\mathrm{tr}\, r(\mathrm{Frob}_v)$ is equal to the eigenvalue of T_v on π_v for all v. This implies (by Cebotarev density) that $r \simeq r_\ell(\pi)$.

Remark 5.4.2. The above theorem is arguably Andrew Wiles' key insight in proving the modularity theorem. It says something much stronger than that suitable ℓ-adic representations are automorphic. It says that at an *integral* level, the ring R^{univ} parametrizing deformations of a mod ℓ Galois representation and the ring $\mathbf{T}_{k,\chi}(U,\mathcal{O})_{\mathfrak{m}}$ parametrizing deformations of a mod ℓ automorphic representation are actually isomorphic.

The next few sections build up to the proof of Theorem 5.4.1.

6. Growth of automorphic forms

In this subsection we will prove Theorem 5.4.1.

6.1. **Setup.**

6.1.1. *Taylor-Wiles primes.* The proof will make use of choices of auxiliary primes.

Let Q be a set of primes with $|Q| < \infty$ and such that Q does not contain v_0 or any prime above ℓ. We further assume that if $v \in Q$, then

- $\overline{r}(\mathrm{Frob}_v)$ has distinct eigenvalues $\alpha_v \neq \beta_v$, and
- $\# k(v) \equiv 1 \pmod{\ell}$.

6.1.2. *Level structure.* We consider groups:

$$U_{0,Q} := \mathrm{GL}_2\left(\widehat{\mathscr{O}_F}^{\{v_0\}\cup Q}\right) \times \mathrm{Iw}^1_{v_0} \times \prod_{v\in Q} \mathrm{Iw}_v$$

$$U_{1,Q} := \mathrm{GL}_2\left(\widehat{\mathscr{O}_F}^{\{v_0\}\cup Q}\right) \times \mathrm{Iw}^1_{v_0} \times \prod_{v\in Q} \mathrm{Iw}^\ell_v$$

Here,

$$\mathrm{Iw}^\ell_v := \left\{ g \in \mathrm{Iw}_v \mid g \mod v \equiv \begin{pmatrix} x & y \\ 0 & z \end{pmatrix}, x/z \text{ has order prime to } \ell \right\}.$$

We have $\mathrm{Iw}^\ell_v \triangleleft \mathrm{Iw}_v$, so $U_{1,Q} \triangleleft U_{0,Q}$. Furthermore, we have:

$$U_{0,Q}/U_{1,Q} =: \Delta_Q = \prod_{v\in Q} \mathrm{Syl}_\ell(\#k(v)^\times).$$

Here, $\mathrm{Syl}_\ell(\#k(v)^\times)$ is the Sylow ℓ-subgroup and is *cyclic*.

6.1.3. Let $S \supseteq Q$ and consider $\mathbf{T} = \mathbf{T}_{k,\chi}(U, \mathscr{O})_{\mathfrak{m}}$. For $v \in Q$, we have:

$$X^2 - T_v X + \#k(v)\chi(\pi_v) \in \mathbf{T}[X]. \tag{6.1.1}$$

By the assumptions in §6.1.1, modulo $\mathfrak{m}$ this polynomial factors as $(X-\alpha_v)(X-\beta_v)$. Now, by Hensel's lemma, we can also factor (6.1.1) as $(X-A_v)(X-B_v)$ for unique $A_v, B_v \in \mathbf{T}$, which will then satisfy

- $A_v + B_v = T_v$,
- $A_v B_v = \#k(v)\chi(\pi_v)$,
- $A_v \equiv \alpha_v \pmod{\mathfrak{m}}$, $B_v \equiv \beta_v \pmod{\mathfrak{m}}$.

6.2. U**-operators.** For $v \in Q$, we define a double coset

$$U_v = U_{0,Q} \begin{pmatrix} \pi_v & 0 \\ 0 & 1 \end{pmatrix} U_{0,Q} = \coprod_{\alpha\in k(v)} \begin{pmatrix} \pi_v & \widetilde{\alpha} \\ 0 & 1 \end{pmatrix} U_{0,Q}$$

where $\widetilde{\alpha}$ is some lift of α. We can also write this with $U_{1,Q}$ instead of $U_{0,Q}$. These are the natural Hecke operators for the $U_{0,Q}$ level structure.

These decompositions of U_v show that the action of U_v on $S_{k,\chi}(U_{0,Q},)$, defined by §3.5.1, is compatible with the action of U_v on $S_{k,\chi}(U_{1,Q},)$ with respect to the inclusion $S_{k,\chi}(U_{0,Q},) \subseteq S_{k,\chi}(U_{1,Q},)$.

We define $\widetilde{\mathbf{U}}_Q := \mathscr{O}[U_v \mid v \in Q]$, and the maximal ideal $\mathfrak{n}_Q := \langle \lambda, U_v - \alpha_v : v \in Q \rangle \subset \widetilde{\mathbf{U}}_Q$. We let $\mathbf{U}_Q$ be the image of $\widetilde{\mathbf{U}}_Q$ in $\mathrm{End}(S_{k,\chi}(U_{0,Q}, \mathscr{O})_{\mathfrak{m}})$, and use the same notation $\mathfrak{n}_Q$ for its image in $\mathbf{U}_Q$.

Lemma 6.2.1. *We have:*

(1) There is an isomorphism

$$S_{k,\chi}(U_{0,Q}, \mathscr{O})^{\oplus 2}_{\mathfrak{m}} \xrightarrow{\sim} S_{k,\chi}(U_{0,Q\cup\{v\}}, \mathscr{O})_{\mathfrak{m}}$$

by the map

$$(\varphi_1, \varphi_2) \mapsto \varphi_1 + \begin{pmatrix} 1 & 0 \\ 0 & \pi_v \end{pmatrix} \varphi_2.$$

(2) There is an isomorphism

$$S_{k,\chi}(U_{0,Q},\mathscr{O})_{\mathfrak{m},\mathfrak{n}_Q} \xrightarrow{\sim} S_{k,\chi}(U_{0,Q\cup\{v\}},\mathscr{O})_{\mathfrak{m},\mathfrak{n}_{Q\cup\{v\}}}$$

by the map

$$\varphi \mapsto -A_v\varphi + \begin{pmatrix} 1 & 0 \\ 0 & \pi_v \end{pmatrix}\varphi.$$

Iterating (2), we get an isomorphism

$$S_{k,\chi}(U,\mathscr{O})_{\mathfrak{m}} \xrightarrow{\sim} S_{k,\chi}(U_{0,Q},\mathscr{O})_{\mathfrak{m},\mathfrak{n}_Q}.$$

This is saying that after a suitable localization, we can treat $U_{0,Q}$ and U as being about the same.

Corollary 6.2.2. *$S_{k,\chi}(U_{1,Q},\mathscr{O})_{\mathfrak{m},\mathfrak{n}_Q}$ is a free $\mathscr{O}[\Delta_Q]$-module, and its Δ_Q-coinvariants are isomorphic to $S_{k,\chi}(U,\mathscr{O})_{\mathfrak{m}}$.*

Proof. We can apply Lemma 3.4.4 since $U_{0,Q}$ is sufficiently small for ℓ. $\square$

Therefore, it suffices to establish (1) and (2).

6.2.1. *Proof of Lemma 6.2.1 (1).* We will show that the same formula in (1) induces isomorphisms:

(a) $S_{k,\chi}(U_{0,Q},\overline{\mathbf{Q}_\ell})_{\mathfrak{m}}^{\oplus 2} \twoheadrightarrow S_{k,\chi}(U_{0,Q\cup\{v\}},\overline{\mathbf{Q}_\ell})_{\mathfrak{m}}$
(b) $S_{k,\chi}(U_{0,Q},\overline{\mathbf{F}}_\ell)_{\mathfrak{m}}^{\oplus 2} \hookrightarrow S_{k,\chi}(U_{0,Q\cup\{v\}},\overline{\mathbf{F}}_\ell)_{\mathfrak{m}}$.

To see that these imply (1), we use the following simple algebra fact: If A,B are finite free $\mathscr{O}$-modules and $f: A\to B$ is a homomorphism such that $f\otimes_{\mathscr{O}}\overline{\mathbf{Q}_\ell}$ is surjective and $f\otimes_{\mathscr{O}}\overline{\mathbf{F}_\ell}$ is injective, then f is an isomorphism. This condition is equivalent to saying that $f\otimes_{\mathscr{O}} L$ is surjective and $f\otimes_{\mathscr{O}}(\mathscr{O}/\lambda)$ is injective. Then, we just apply the facts (cf. §3.4.1) that $S_{k,\chi}(U_{0,Q},\mathscr{O})\otimes_{\mathscr{O}}\overline{\mathbf{Q}}_\ell \xrightarrow{\sim} S_{k,\chi}(U_{0,Q},\overline{\mathbf{Q}_\ell})$ (which is true simply because $\overline{\mathbf{Q}}_\ell$ is $\mathbf{Z}_\ell$-flat) and $S_{k,\chi}(U_{0,Q},\mathscr{O})\otimes_{\mathscr{O}}\overline{\mathbf{F}}_\ell \xrightarrow{\sim} S_{k,\chi}(U_{0,Q},\overline{\mathbf{F}}_\ell)$ because $U_{0,Q}$ is sufficiently small for ℓ.

Now, we prove (a). We change coefficients to $\mathbf{C}$ for convenience. We have

$$S_{k,\chi}(U_{0,Q},\mathbf{C})_{\mathfrak{m}} \cong \bigoplus_{\pi:\,\overline{r}_\ell(\pi)\cong\overline{r}} (\pi^\infty)^{U_{0,Q}} \cong \bigoplus_{\pi:\,\overline{r}_\ell(\pi)\cong\overline{r}} (\pi^{\infty,v})^{U_{0,Q}^v}\otimes\pi_v^{\mathrm{GL}_2(\mathscr{O}_{F,v})}.$$

(Here we used that since $\overline{r}$ is irreducible, the components corresponding to $\phi\in\mathscr{C}_{k,\chi}$ vanish upon localization at $\mathfrak{m}$). Similarly,

$$S_{k,\chi}(U_{1,Q},\mathbf{C})_{\mathfrak{m}} \cong \bigoplus_{\pi:\,\overline{r}_\ell(\pi)\cong\overline{r}} (\pi^\infty)^{U_{1,Q\cup\{v\}}} \cong \bigoplus_{\pi:\,\overline{r}_\ell(\pi)\cong\overline{r}} (\pi^{\infty,v})^{U_{1,Q}^v}\otimes\pi_v^{\mathrm{Iw}_v}.$$

It suffices to prove that

$$(\pi_v^{\mathrm{GL}_2(\mathscr{O}_{F,v})})^{\oplus 2} \twoheadrightarrow \pi_v^{\mathrm{Iw}_v}$$

under the map

$$(\phi_1,\phi_2)\mapsto \phi_1 + \begin{pmatrix} 1 & \\ & \pi_v \end{pmatrix}\phi_2.$$

We'll do this by explicitly analyzing representations with Iwahori fixed vector.

Fix some such π. There are two possibilities: π_v is either ramified or unramified. If π_v is ramified, then

$$\operatorname{rec}(\pi_v) \simeq \left(\begin{pmatrix} \epsilon_\ell \phi & 0 \\ 0 & \phi \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right)$$

with ϕ an unramified character. Thus, $r_\ell(\pi)(\widetilde{\mathrm{Frob}_v})$ has eigenvalues α and $\alpha \# k(v)$ for any lift $\widetilde{\mathrm{Frob}_v} \in G_{F_v}$ of Frob_v. But $\overline{r_\ell(\pi)} = \overline{r}$, which says that $\overline{r}(\mathrm{Frob}_v)$ has eigenvalues $\overline{\alpha}$ and $\overline{\alpha \# k(v)} = \overline{\alpha}$ which contradicts our assumptions on v in §6.1.1.

If π_v is unramified, then it has the form

$$\pi_v \simeq \operatorname{Ind}_{B(F_v)}^{G(F_v)}(\psi_1, \psi_2)$$

with ψ_1, ψ_2 unramified characters such that $\psi_1/\psi_2 \not\simeq |\cdot|_v^{\pm 1}$. Using the Cartan decomposition,

$$\mathrm{GL}_2(F_v) = B(F_v)\,\mathrm{GL}_2(\mathscr{O}_{F,v})$$

we find an isomorphism

$$\pi_v^{\mathrm{GL}_2(\mathscr{O}_{F,v})} \xrightarrow{\sim} \mathbf{C}$$

by the map sending ϕ to $\phi(\mathrm{Id})$.

Now using also the Bruhat decomposition, we get

$$\mathrm{GL}_2(F_v) = B(F_v)\mathrm{Iw}_v \coprod B(F_v) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \mathrm{Iw}_v$$

which shows that

$$\pi_v^{\mathrm{Iw}_v} \xrightarrow{\sim} \mathbf{C}^2$$

by the map sending

$$\phi \mapsto (\phi(1), \phi(w)).$$

Now we examine the map

$$\begin{tikzcd}
\left(\operatorname{Ind}_{B(F_v)}^{G(F_v)}(\psi_1,\psi_2)^{\mathrm{GL}_2(\mathscr{O}_{F,v})}\right)^{\oplus 2} \arrow[r] \arrow[d, equal] & \operatorname{Ind}_{B(F_v)}^{\mathrm{GL}_2(F_v)}(\psi_1,\psi_2)^{\mathrm{Iw}_v} \arrow[d, equal] \\
\mathbf{C}^2 \arrow[r, dashed] & \mathbf{C}^2
\end{tikzcd}$$

It is given by

$$\begin{tikzcd}
(\phi_1, \phi_2) \arrow[r] \arrow[d] & \phi_1 + \begin{pmatrix} 1 & \\ & \pi_v \end{pmatrix} \phi_2 \arrow[d] \\
(\phi_1(\mathrm{Id}), \phi_1(\mathrm{Id})) \arrow[r, dashed] & \phi_1(\mathrm{Id}) + \begin{pmatrix} 1 & \\ & \pi_v \end{pmatrix} \phi_2(\mathrm{Id}), \phi_1(w) + \begin{pmatrix} 1 & \\ & \pi_v \end{pmatrix} \phi_2(w) \arrow[d, equal] \\
& (\phi_1(\mathrm{Id}) + \#k(v)^{1/2}\psi_2(v)\cdot\phi_2(1), \phi_1(1) + \#k(v)^{-1/2}\psi_1(v)\cdot\phi_2(\mathrm{Id}))
\end{tikzcd}$$

In other words, the map is given on $\mathbf{C}^2$ by applying

$$\begin{pmatrix} 1 & \#k(v)^{1/2}\psi_2(v) \\ 1 & \#k(v)^{-1/2}\psi_1(v) \end{pmatrix} \tag{6.2.1}$$

and by the assumption of unramifiedness, this is invertible.

This completes the proof of (a). Next we turn to (b). Suppose you have (x_1, x_2) in the kernel, so

$$x_1 = -\begin{pmatrix} 1 & \\ & \pi_v \end{pmatrix} x_2.$$

Let $M \subset S_{k,\chi}(\overline{\mathbf{F}}_\ell)$ be the $\mathrm{GL}_2(F_v)$-submodule generated by x_1, and N the maximal $\mathrm{GL}_2(F_v)$-submodule not containing x_1. Then we have that (x_1, x_2) is in the kernel of the induced map

$$\left((M/N)^{\mathrm{GL}_2(\mathcal{O}_{F,v})}\right)^{\oplus 2} \to (M/N)^{\mathrm{Iw}_v}.$$

Now, by the classification of unramified represents, we have two possibilities for (M/N):

(i) $M/N \cong \psi_0 \circ \det$,

(ii) $M/N \cong \mathrm{Ind}_{B(F_v)}^{\mathrm{GL}_2(F_v)}(\psi_1, \psi_2)$ with ψ_1, ψ_2 unramified and $\psi_1 \neq \psi_2$, which explicitly is

$$\left\{\phi \colon \mathrm{GL}_2(F_v) \to \overline{\mathbf{F}}_\ell \colon \phi\left(\begin{pmatrix} x & y \\ & z \end{pmatrix} g\right) = \psi_1(x)\psi_2(y)\phi(g)\right\}.$$

(Since $\#k(v) \equiv 1 \pmod{\ell}$, we can drop the normalizing factor.)

We'll analyze these in turn. In case (ii), the same computation as before shows that the map is given by (6.2.1), which since $\#k(v) \equiv 1 \pmod{\ell}$ and $\psi_1 \neq \psi_2$ is invertible. In case (i), we have $M/N \cong \overline{\mathbf{F}}_\ell$ and the fact that its $\mathfrak{m}$-adic localization is non-zero shows that the Hecke action is given by

$$T_v x = (\#k(v) + 1)\psi(\pi_v)x.$$

It also tells us that $\chi(\pi_v) = \psi^2(\pi_v)$. Hence the Hecke eigenvalues α_v, β_v satisfy

$$\alpha_v + \beta_v = (\#k(v) + 1)\psi(\pi_v) = 2\psi(\pi_v)$$

and

$$\alpha_v \beta_v = \#k(v)\chi(\pi_v) = \psi(\pi_v)^2.$$

So α_v, β_v are the roots of

$$x^2 - 2\psi(\pi_v)x + \psi(\pi_v)^2 = (x - \psi(\pi_v))^2$$

which implies that $\alpha_v \equiv \beta_v$, contradicting our assumptions in §6.1.1.

In case (ii),

6.2.2. *Proof of Lemma 6.2.1 (2).* Consider the diagram

$$\begin{array}{ccc} S_{k,\chi}(U_{0,Q}; \mathcal{O})^{\oplus 2} & \longrightarrow & S_{k,\chi}(U_{0,Q \cup \{v\}}; \mathcal{O}) \\ & & \big\downarrow U_v \\ S_{k,\chi}(U_{0,Q}; \mathcal{O})^{\oplus 2} & \longrightarrow & S_{k,\chi}(U_{0,Q \cup \{v\}}; \mathcal{O}) \end{array}$$

We claim that we can fill this in to a commutative diagram

$$\begin{tikzcd} S_{k,\chi}(U_{0,Q};\mathscr{O})^{\oplus 2} \arrow[r] \arrow[d, dashed] & S_{k,\chi}(U_{0,Q\cup\{v\}};\mathscr{O}) \arrow[d, "U_v"] \\ S_{k,\chi}(U_{0,Q};\mathscr{O})^{\oplus 2} \arrow[r] & S_{k,\chi}(U_{0,Q\cup\{v\}};\mathscr{O}) \end{tikzcd}$$

with the dashed map being

$$\begin{pmatrix} T_v & \#k(v)\chi(\pi_v) \\ -1 & 0 \end{pmatrix} \tag{6.2.2}$$

Proof of claim. We can just check that the claimed formula works. We apply U_v to the image of (φ_1, φ_2):

$$\begin{aligned} \varphi_1 + \begin{pmatrix} 1 & \\ & \pi_v \end{pmatrix}\varphi_2 \overset{U_v}{\mapsto} & \sum_{\alpha\in k(v)} \begin{pmatrix} \pi_v & \alpha \\ & 1 \end{pmatrix}\varphi_1 + \sum_{\alpha\in k(v)} \begin{pmatrix} \pi_v & \alpha \\ & 1 \end{pmatrix}\begin{pmatrix} 1 & \\ & \pi_v \end{pmatrix}\varphi_2 \\ &= \sum_{\alpha\in k(v)} \begin{pmatrix} \pi_v & \alpha \\ & 1 \end{pmatrix}\varphi_1 + \sum_{\alpha\in k(v)} \begin{pmatrix} \pi_v & \alpha\pi_v \\ & \pi_v \end{pmatrix}\varphi_2 \\ &= T_v\varphi_1 - \begin{pmatrix} 1 & \\ & \pi_v \end{pmatrix}\varphi_1 + \#k(v)\chi(\pi_v)\varphi_2. \end{aligned}$$

□

So U_v acts on $S_{k,\chi}(U_{0,Q};\mathscr{O})^{\oplus 2}_{\mathfrak{m},\mathfrak{n}_Q} \xrightarrow{\sim} S_{k,\chi}(U_{0,Q\cup\{v\}};\mathscr{O})_{\mathfrak{n},\mathfrak{n}_Q}$ by

$$\begin{pmatrix} A_v + B_v & A_v B_v \\ -1 & 0 \end{pmatrix}. \tag{6.2.3}$$

By splitting into the eigenspaces of the linear transformation (6.2.3), we get a decomposition

$$S_{k,\chi}(U_{0,Q};\mathscr{O})^{\oplus 2}_{\mathfrak{m},\mathfrak{n}_Q} \cong \left\{\begin{pmatrix} A_v\chi \\ -\chi \end{pmatrix} : x \in S_{k,\chi}(U_{0,Q};\mathscr{O})_{\mathfrak{m},\mathfrak{n}_Q}\right\} \oplus \left\{\begin{pmatrix} B_v x \\ -x \end{pmatrix} : x \in S_{k,\chi}(U_{0,Q};\mathscr{O})_{\mathfrak{m},\mathfrak{n}_Q}\right\} \tag{6.2.4}$$

You have to be careful about direct sums because we are working integrally, but the point is that

$$\begin{pmatrix} A_v & B_v \\ -1 & -1 \end{pmatrix}$$

is invertible under our assumptions, which imply that $A_v - B_v$ reduces to a unit in $\mathbf{F}$.

Localizing at $(\mathfrak{m}, U_v - A_v) \subset \mathbf{T}$, we get that $U_v - B_v$ is a unit since $A_v - B_v$ reduces to a unit. So the second summand, which is the $U_v = B_v$ eigenspace, disappears in the localization. Conversely, if $f(U_v) \notin (\mathfrak{m}, U_v - A_v)$ then writing $f(X) = (X - A_v)q(X) + C_v$ with $C_v \notin \mathfrak{m}$, we find that $f(U_v)$ acts as C_v, which is a unit. Hence the localization does nothing to the first summand, and we conclude that

$$S_{k,\chi}(U_{0,Q\cup\{v\}};\mathscr{O})_{\mathfrak{n},\mathfrak{n}_{Q\cup\{v\}}} \xrightarrow{\sim} (S_{k,\chi}(U_{0,Q};\mathscr{O})^{\oplus 2}_{\mathfrak{m},\mathfrak{n}_Q})_{\mathfrak{n}_v} \cong S_{k,\chi}(U_{0,Q};\mathscr{O})_{\mathfrak{m},\mathfrak{n}_Q}.$$

We have proved part (2).

6.3. **Adding level structure.** Recall that we deduced from Lemma 6.2.1, by inductive application of (2), that

$$S_{k,\chi}(U;\mathcal{O})_{\mathfrak{m}} \cong S_{k,\chi}(U_{0,Q};\mathcal{O})_{\mathfrak{m},\mathfrak{n}_Q},$$

and that $S_{k,\chi}(U_{1,Q};\mathcal{O})_{\mathfrak{m},\mathfrak{n}_Q}$ is a finite free $\mathcal{O}[\Delta_Q]$-module, which implies

$$(S_{k,\chi}(U_{1,Q};\mathcal{O})_{\mathfrak{m},\mathfrak{n}_Q})_{\Delta_Q} \cong S_{k,\chi}(U;\mathcal{O})_{\mathfrak{m}}.$$

6.3.1. We're now going to piece together the key picture in the patching argument. We have $\mathbf{T} = \mathbf{T}_{k,\chi}(U;\mathcal{O})_{\mathfrak{m}}$ acting on $S = S_{k,\chi}(U;\mathcal{O})_{\mathfrak{m}}$. We have Δ_Q acting on $S_Q = S_{k,\chi}(U_{1,Q};\mathcal{O})_{\mathfrak{m},\mathfrak{n}_Q} \twoheadrightarrow S$, which factors through an isomorphism

$$\begin{tikzcd}
 & S_Q \arrow[dl] \arrow[d, two heads] \\
(S_Q)_{\Delta_Q} \arrow[r, dashed, "\sim"] & S
\end{tikzcd}$$

There's the Hecke algebra $\mathbf{T}_Q$ acting on S_Q (the image of $\mathbf{T}_{k,\chi}(U_{1,Q};\mathcal{O})_{\mathfrak{m}}$ in the localized $\mathrm{End}_{\mathcal{O}}(S_{k,\chi}(U_{1,Q};\mathcal{O})_{\mathfrak{m},\mathfrak{n}_Q})$. Also we have a quotient map $\mathbf{T}_Q \twoheadrightarrow \mathbf{T}$, induced by the action on $S \hookrightarrow S_Q$.

$$\begin{tikzcd}
 & S_Q \arrow[dl] \arrow[d, two heads] & S_Q & \curvearrowright & \mathbf{T}_Q \arrow[d, two heads] \\
(S_Q)_{\Delta_Q} \arrow[r, dashed, "\sim"] & S & S \arrow[u, hook] & \curvearrowright & \mathbf{T}
\end{tikzcd}$$

6.3.2. We also have the universal deformation ring $R^{\mathrm{univ}}_{\{v_0\}} = R^{\mathrm{univ}}$, and in §5.4 we produced a map

$$R^{\mathrm{univ}}_{\{v_0\}} = R^{\mathrm{univ}} \twoheadrightarrow \mathbf{T} = \mathbf{T}_{k,\chi}(U;\mathcal{O})_{\mathfrak{m}}.$$

(The map is surjective because it hits the generators of $\mathbf{T}$: in fact, $\mathrm{Tr}(\mathrm{Frob}_v) \mapsto T_v$.)

The pair $(R^{\mathrm{univ}}_{Q\cup\{v_0\}}, r^{\mathrm{univ}}_{Q\cup\{v_0\}}) = (R^{\mathrm{univ}}_Q, r^{\mathrm{univ}}_Q)$ represents the functor on complete noetherian local $\mathcal{O}$-algeras with residue field $\mathbf{F}$, sending A to continuous representations $\rho : G_F \to \mathrm{GL}_2(A)$ such that $\rho \pmod{\mathfrak{m}_A} = \overline{r}$, ρ unramified away from $1, v_0, Q$ and ρ is FL above ℓ, modulo equivalence.

6.3.3. The key picture is the diagram

$$\begin{tikzcd}
 & S_Q \arrow[dl] \arrow[d, two heads] & S_Q & \curvearrowright & \mathbf{T}_Q \arrow[d, two heads] & R^{\mathrm{univ}}_Q \arrow[l, two heads] \arrow[d, two heads] \\
(S_Q)_{\Delta_Q} \arrow[r, dashed, "\sim"] & S & S \arrow[u, hook] & \curvearrowright & \mathbf{T} & R^{\mathrm{univ}} \arrow[l, two heads]
\end{tikzcd}$$

Lemma 6.3.1. *We have the following.*

(1) If $v \in Q$, then

$$r^{\mathrm{univ}}_Q|_{G_{F_v}} \sim \begin{pmatrix} \psi_\alpha & \\ & \psi_\beta \end{pmatrix}$$

where

$$\psi_\alpha(\mathrm{Frob}_v) \equiv \alpha_v \pmod{\mathfrak{m}},$$
$$\psi_\beta(\mathrm{Frob}_v) \equiv \beta_v \pmod{\mathfrak{m}}.$$

(2) *If $v \in Q$, then*

$$\begin{tikzcd} \mathcal{O}_{F_v}^\times \arrow[r] \arrow[d, "\mathrm{Art}_{F_v}"] & k(v)^\times \arrow[r] & \Delta_v \arrow[dl, dashed] \arrow[d, dashed] \\ G_{F_v}^{\mathrm{ab}} \arrow[r, "\psi_\alpha"] & (R_q^{\mathrm{univ}})^\times \arrow[r] & \mathbf{T}_Q^\times \end{tikzcd}$$

and the resulting action of Δ_v agrees with that of the diamond operators.

In other words, we have $\mathcal{O}[\Delta_Q] \to R_Q^{\mathrm{univ}}$ which induces the map $\Delta_Q \to (R_Q^{\mathrm{univ}})^\times$. Also, if $\mathfrak{a}_Q = \langle \delta - 1, \delta \in \Delta_Q \rangle$ then $R_Q^{\mathrm{univ}}/\mathfrak{a}_Q = R^{\mathrm{univ}}$. (Indeed, this quotient imposes that inertia acts trivially, so it's just saying that the representation is unramified at Q.)

Proof. (1) Let $\widetilde{\mathrm{Frob}}_v \in G_{F_v}$ be a lift of $\mathrm{Frob}_v \in G_{k(v)}$. Consider the characteristic polynomial of this lift. Modulo $\mathfrak{m}$, this factors as $(X-\alpha_v)(X-\beta_v)$, so since R_Q^{univ} is complete we can apply Hensel's Lemma to get a factorization

$$\mathrm{Char}_{r_Q^{\mathrm{univ}}}(\widetilde{\mathrm{Frob}}_v)(X) = (X-A)(X-B)$$

with $A, B \in R_Q^{\mathrm{univ}}$ reducing to α_v, β_v. Because $A - B \notin \mathfrak{m}_{R_Q^{\mathrm{univ}}}$, it is a unit, and we can diagonalize this matrix. So we can assume that

$$r_Q^{\mathrm{univ}}(\widetilde{\mathrm{Frob}}_v) = \begin{pmatrix} A & \\ & B \end{pmatrix}.$$

What we need is that $r_Q^{\mathrm{univ}}(I_{F_V})$ is also diagonal. Since this acts trivially residually, it is a pro-ℓ group, hence factors through tame inertia. Let σ be a topological generator; we need to show that $r_Q^{\mathrm{univ}}(\sigma)$ is diagonal. We know that

$$r_Q^{\mathrm{univ}}(\widetilde{\mathrm{Frob}}_v \sigma \widetilde{\mathrm{Frob}}_v^{-1}) = r_Q^{\mathrm{univ}}(\sigma^{\# k(v)}).$$

Let $I \subset R_Q^{\mathrm{univ}}$ be the ideal generated by the off-diagonal elements of $r_Q^{\mathrm{univ}}(\sigma)$. We want to show that $I = 0$. It suffices to prove that $\mathfrak{m} I = I$. We'll do this by showing that

$$r_Q^{\mathrm{univ}}(\sigma) \pmod{\mathfrak{m} I} =: \begin{pmatrix} x & y \\ z & w \end{pmatrix}$$

also has $z = y = 0$ in $R_Q^{\mathrm{univ}}/\mathfrak{m} I$. Using the relation, we find that

$$\begin{pmatrix} x & (B/A)y \\ (A/B)z & w \end{pmatrix} = \left(\begin{pmatrix} x & \\ & w \end{pmatrix} + \begin{pmatrix} 0 & y \\ z & 0 \end{pmatrix} \right)^q = \begin{pmatrix} x^q & \\ & y^q \end{pmatrix} + q \begin{pmatrix} 0 & y \\ z & 0 \end{pmatrix}.$$

The induction assumption allows us to assume that $y, z \in I$, and since we are working modulo $\mathfrak{m}$ we can replace A, B with α_v, β_v and x, w with 1 (which are their reductions modulo $\mathfrak{m}$). So we get $(\beta_v/\alpha_v - q)y = 0$ and $(\beta_v/\alpha_v - q)z = 0$, but this forces $y = z = 0$ because

$$(\beta_v/\alpha_v - q) \equiv \beta_v/\alpha_v - 1 \neq 0 \pmod{\mathfrak{m}}.$$

(2) We check that the two actions of Δ_Q on S_Q agree. The more obvious action just comes from the fact that $S_Q = S_{k,\chi}(U_{1,Q};\mathcal{O})_{\mathfrak{m};\mathfrak{n}_q}$ and $U_{0,Q}/U_{1,Q}$ is Δ_Q; the other comes from the Galois action.

It suffices to check this after tensoring up to $\mathbf{C}$,

$$S_Q = \bigoplus_{\pi:\,\overline{r}_\ell(\pi)\cong r} (\pi^\infty \otimes ||\det||^{1/2})^{U_{1,Q}}_{\mathfrak{n}_Q}.$$

For $v \in Q$, we may assume that $\pi_v^{\mathrm{Iw}^1_v} \neq 0$, otherwise it doesn't contribute to the direct sum. Hence, by the classification of such representations, we have that π_v is a subquotient of

$$\mathrm{Ind}_{B(F_v)}^{G(F_v)}(\theta_\alpha \times \theta_\beta).$$

Then the semisimplification of the local Galois representation has the form

$$(r_\ell(\pi_v)|_{W_{F_v}})^{\mathrm{ss}} = \theta_\alpha \circ \mathrm{Art}_{F_v}^{-1} \oplus \theta_\beta \circ \mathrm{Art}_{F_v}^{-1}$$

for tamely ramified $\theta_\alpha, \theta_\beta$. The reduction is unramified, and we can choose the labeling so that $\theta_\alpha(\pi_v) \equiv \alpha_v \pmod{\lambda}$ and $\theta_\beta(\pi_v) \equiv \beta_v \pmod{\lambda}$.

The action of $\mathcal{O}[\Delta_v]$ on S_Q is induced by the composition $\mathcal{O}[\Delta_v] \to R_Q^{\mathrm{univ}} \to \mathbf{T}_Q$ comes from θ_α. We have to compare this with the action of $U_{0,Q}/U_{1,Q}$ on

$$\mathrm{Ind}_{B(F_v)}^{\mathrm{GL}_2(F_v)}(\theta_\alpha \times \theta_\beta \otimes ||\det||^{1/2})^{\mathrm{Iw}^1} \xrightarrow{\sim} \mathbf{C}^2 \tag{6.3.1}$$

where the identification is via

$$\varphi \mapsto (\varphi(\mathrm{Id}), \varphi(w)).$$

How does Δ_v act? A typical element of Δ_v is

$$\begin{pmatrix} \delta & \\ & 1 \end{pmatrix}, \quad \delta \in \mathcal{O}_{F,v}^\times.$$

Then

$$\begin{pmatrix} \delta & \\ & 1 \end{pmatrix} \cdot \varphi(\mathrm{Id}) = \varphi\left(\begin{pmatrix} \delta & \\ & 1 \end{pmatrix}\right) = \theta_\alpha(\delta)\varphi(\mathrm{Id})$$

and

$$\begin{pmatrix} \delta & \\ & 1 \end{pmatrix} \cdot \varphi(w) = \varphi\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \delta & \\ & 1 \end{pmatrix}\right) = \varphi\left(\begin{pmatrix} 1 & \\ & \delta \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right) = \theta_\beta(\delta)\varphi(w).$$

So we found that δ acts as on (6.3.1) by the matrix

$$\begin{pmatrix} \theta_\alpha(\delta) & \\ & \theta_\beta(\delta) \end{pmatrix}$$

with respect to the coordinates given by evaluation at Id and w. We have to show that upon localization at $\mathfrak{n}_Q$, it acts through θ_α.

In conclusion, what we've computed is that

$$\begin{aligned}(U_v\cdot\varphi)(\mathrm{Id}) &= \sum_{\gamma\in k(v)}\varphi\begin{pmatrix}\pi_v & \gamma\\ 0 & 1\end{pmatrix}\#k(v)^{-1/2}\\ &= \#k(v)\theta_\alpha(\pi)\cdot\#k(v)^{-1/2}\cdot\#k(v)^{-1/2}\cdot\varphi(\mathrm{Id})\\ &= \theta_\alpha(\pi_v)\varphi(\mathrm{Id})\end{aligned}$$

and

$$\begin{aligned}(U_v\cdot\varphi)(w) &= \sum_{\gamma\in k(v)}\varphi\begin{pmatrix}0 & 1\\ \pi_v & \tilde{\gamma}\end{pmatrix}\#k(v)^{-1/2}\\ &= \varphi\left(\begin{pmatrix}1 & \\ & \pi_v\end{pmatrix}w\right)\#k(v)^{-1/2}+\sum_{\gamma\in k(v)-0}\varphi\left(\begin{pmatrix}-\gamma^{-1}\pi_v & 1\\ 0 & \gamma\end{pmatrix}\begin{pmatrix}1 & 0\\ \pi_v\gamma^{-1} & 1\end{pmatrix}\right)\#k(v)^{-1/2}\\ &= \theta_\beta(\pi_v)\varphi(w)+\sum_{\gamma\in k(v)-0}\theta_\alpha(-\gamma^{-1}\pi_v)\theta_\beta(\gamma)\#k(v)^{-1}\varphi(\mathrm{Id})\\ &= \theta_\beta(\pi_v)\varphi(w)+\theta_\alpha(-\pi_v)\#k(v)^{-1}\sum_{\gamma\in k(v)^\times}\frac{\theta_\beta}{\theta_\alpha}(\gamma)\varphi(\mathrm{Id}).\end{aligned}$$

Hence U_v acts as

$$\begin{pmatrix}\theta_\alpha(\pi_v) & 0\\ \theta_\alpha(-\pi_v)\#k(v)^{-1}\sum_{\gamma\in k(v)^\times}(\theta_\beta/\theta_\alpha)(\gamma) & \theta_\beta(\pi_v)\end{pmatrix}$$

If $\theta_\beta/\theta_\alpha|_{\Delta_v}=1$ then we win, as then Δ_v acts on the invariants by $\theta_\alpha(\delta)$ because $\theta_\alpha=\theta_\beta$. Otherwise, we sum a non-trivial character and this bottom left entry becomes 0, so U_v acts as

$$\begin{pmatrix}\theta_\alpha(\pi_v) & 0\\ 0 & \theta_\beta(\pi_v)\end{pmatrix}.$$

Then $U_v-\theta_\beta(\pi_v)$, which is not in $\mathfrak{n}_Q$ hence acts as a unit in the localization, kills the second factor and we find that $\delta\in\Delta_v$ acts on $(\mathbf{C}^{\oplus 2})_{\mathfrak{n}_Q}$ by $\theta_\alpha(\delta)$. □

6.3.4. Let's go back to the big picture:

$$\begin{tikzcd} S_Q \arrow[loop above, phantom, ""] & \curvearrowright & \mathbf{T}_Q \arrow[d, two heads] & R_Q^{\mathrm{univ}} \arrow[l, two heads] \arrow[d, two heads]\\ S \arrow[u, hook] & \curvearrowright & \mathbf{T} & R^{\mathrm{univ}} \arrow[l, two heads]\end{tikzcd}$$

This saying that we pick a finite set of favorable ("Taylor-Wiles") primes, and we simultaneously consider a Hecke algebra at the corresponding level and a Galois deformation ring allowing ramification at these primes. This equips R_Q^{univ} with the structure of an $\mathscr{O}[\Delta_Q]$-algebra.

We now know that S_Q is finite free over $\mathscr{O}[\Delta_Q]$, and for $\mathfrak{a}_Q\subset\mathscr{O}[\Delta_Q]$ the augmentation ideal, the trace map induces an isomorphism

$$\mathrm{Tr}\colon S_{Q,\mathfrak{a}_Q}\xrightarrow{\sim}S.$$

We want to prove that $R^{\mathrm{univ}} \xrightarrow{\sim} \mathbf{T}$. We don't need to know anything more about automorphic forms, other than this paragraph. The moral is that "when you relax the ramification, you see that S is as large as it could be".

7. Galois deformation rings

7.1. **Generators.** We need to know something about the size of the rings R^{univ} and R_Q^{univ}.

The universal deformation ring R_Q^{univ} is a complete local $\mathcal{O}$-algebra, say with maximal ideal $\mathfrak{m}_Q$.

Lemma 7.1.1. *If R is a complete Noetherian ring over $\mathcal{O}$, with maximal ideal $\mathfrak{m}$, then it is topologically generated over $\mathcal{O}$ by $\dim_{\mathbf{F}}(\mathfrak{m}/(\lambda, \mathfrak{m}^2))$ elements.*

Proof. Pick a basis $r_1, \dots, r_n$ and lift them to $\tilde{r}_i \in \mathfrak{m}$. The choice of lifts defines a map

$$\mathcal{O}[[x_1, \dots, x_n]] \to R.$$

Since $R = \varprojlim R/\mathfrak{m}^N$, it suffices to prove that

$$\mathcal{O}[[x_1, \dots, x_n]] \to R/\mathfrak{m}^N.$$

We induct on N. We can approximate up to $\mathfrak{m}^{N-1}$ by induction. So it suffices to show that

$$(x_1, \dots, x_n)^N \twoheadrightarrow \mathfrak{m}^N/\mathfrak{m}^{N+1}.$$

By Nakayama, it suffices to show that

$$(x_1, \dots, x_n)^N \twoheadrightarrow \mathfrak{m}^N/(\lambda, \mathfrak{m}^{N+1}).$$

The induction hypothesis gives that

$$(x_1, \dots, x_n)^{N-1} \twoheadrightarrow \mathfrak{m}^{N-1}/\mathfrak{m}^N.$$

and then we consider the diagram

$$\begin{tikzcd}
(x_1, \dots, x_n)^{N-1} \otimes (x_1, \dots, x_n) \arrow[r, two heads] \arrow[d, two heads] & \mathfrak{m}^{N-1}/\mathfrak{m}^N \otimes (\mathfrak{m}/(\lambda, \mathfrak{m}^2)) \arrow[d, two heads] \\
(x_1, \dots, x_n)^N \arrow[r] & \mathfrak{m}^N/(\lambda, \mathfrak{m}^{N+1})
\end{tikzcd}$$

□

7.2. **The tangent space.** We now analyze the space $(\mathfrak{m}_Q/(\lambda, \mathfrak{m}_Q^2))$.

Lemma 7.2.1. *We have*

$$\mathrm{Hom}_{\mathbf{F}}(\mathfrak{m}_Q, (\lambda, \mathfrak{m}_Q^2), \mathbf{F}) \cong \mathrm{Hom}_{\mathcal{O}}(R_Q^{\mathrm{univ}}, \mathbf{F}[\epsilon]/\epsilon^2)$$

where the right hand side denotes local $\mathcal{O}$-algebra homomorphisms.

Proof. For $\alpha \in \mathrm{Hom}_{\mathscr{O}}(R_Q^{\mathrm{univ}}, \mathbf{F}[\epsilon]/\epsilon^2)$, consider $\alpha|_{\mathfrak{m}_Q}$. Since α is a local homomorphism, $\alpha|_{\mathfrak{m}_Q}$ lands in $\epsilon\mathbf{F}$, and we associate to it $\frac{1}{\epsilon}\alpha|_{\mathfrak{m}_Q} \in \mathrm{Hom}_{\mathbf{F}}(\mathfrak{m}_Q, (\lambda, \mathfrak{m}_Q^2), \mathbf{F})$.

Conversely, we can send β to the homomorphism $R_Q^{\mathrm{univ}} \to \mathbf{F}[\epsilon]/\epsilon^2$ taking

$$\underbrace{a}_{\in\mathscr{O}} + \underbrace{b}_{\in\mathfrak{m}_Q} \mapsto a \pmod{\lambda} + \beta(b)\epsilon.$$

□

7.2.1. *Interpretation as deformations.* The latter space $\mathrm{Hom}_{\mathscr{O}}(R_Q^{\mathrm{univ}}, \mathbf{F}[\epsilon]/\epsilon^2)$ has an interpretation in terms of deformations, by the universal property: it's a subspace of the deformations

$$G_F \to \mathrm{GL}_2(\mathbf{F}[\epsilon]/\epsilon^2)$$

of the form $\sigma \mapsto (1+\phi(\sigma)\epsilon)\overline{r}(\sigma)$. The condition for this to be a homomorphism is that

$$(1+\phi(\sigma\tau)\epsilon)\overline{r}(\sigma\tau) = (1+\phi(\sigma)\epsilon)\overline{r}(\sigma)(1+\phi(\tau)\epsilon)\overline{r}(\tau).$$

Exercise 7.2.2. Check that this is equivalent to asking that

$$\phi(\sigma\tau) = \phi(\sigma) + \mathrm{ad}\,\overline{r}(\sigma)\phi(\tau),$$

i.e. that $\phi \in Z^1(G_F; \mathrm{ad}\,\overline{r})$.

We are now going to digest in these terms the other conditions that we imposed in our deformations.

7.2.2. *Fixed determinant.* In our deformation problem we fixed the determinants of the representations. This is equivalent to imposing

$$\det(1+\phi(\sigma)\epsilon) = 1,$$

which amounts to $\mathrm{Tr}\,\mathrm{ad}(\sigma) = 0$. We let $\mathrm{ad}^0(\overline{r}) \subset \mathrm{ad}(\overline{r})$ be the subspace of trace 0 matrices.

7.2.3. *Controlling ramification.* We need the deformation to be unramified away from Q and ℓ. To encode this, we introduce some notation. Let F_T be the maximal extension of F unramified outside T, and we let $G_{F,T} = \mathrm{Gal}(F_T/F)$. Then we are saying that ϕ factors through $G_{F,Q\cup\{\ell\}}$.

7.2.4. *Fontaine-Laffaille.* We imposed that the deformations be Fontaine-Laffaille. We'll explicate this condition shortly.

7.2.5. *Equivalence relation.* We imposed an equivalence relation on lifts, declaring them to be equivalent if they are conjugate by a matrix reducing to the identity over $\mathbf{F}$.

In the case at hand, this means we mod out by conjugation by things of the form $1 + a\epsilon$ for $a \in M_{2\times 2}(\mathbf{F})$.

Exercise 7.2.3. Check that this conjugation takes ϕ to the function

$$\sigma \mapsto a + \phi(\sigma) - \mathrm{ad}\,\overline{r}(\sigma)a.$$

This amounts to the usual equivalence relation in Galois cohomology.

Hence, the space $\mathrm{Hom}_{\mathcal{O}}(R_Q^{\mathrm{univ}}, \mathbf{F}[\epsilon]/\epsilon^2)$ we're interested in can be written as

$$\begin{aligned}\{\phi \in Z^1(G_{F,Q\cup\{\ell\}}, \mathrm{ad}^0\overline{r})\colon (1+\phi\epsilon)\overline{r}|_{G_{F_v}} \text{ is FL at all } v \mid \ell\}/(\sigma \mapsto a+\phi(\sigma) - \mathrm{ad}\,\overline{r}(\sigma)a)\\ = \{[\phi] \in H^1(G_{F,Q\cup\{\ell\}}, \mathrm{ad}^0\overline{r})\colon \mathrm{Res}[\phi]|_{G_{F_v}} \text{ is FL at all } v \mid \ell\}.\end{aligned} \tag{7.2.1}$$

7.2.6. *Unraveling the FL condition.* We want to explicate the FL condition.

Lemma 7.2.4. *For all $v \mid \ell$, we have*

$$H^1(G_{F_v}, \mathrm{ad}\,\overline{r}) \cong \mathrm{Ext}^1_{\mathbf{F}[G_{F_v}]}(\overline{r}, \overline{r}). \tag{7.2.2}$$

Proof. Indeed, a cocycle $(1+\phi\epsilon)\overline{r}$ is manifestly an extension of $\overline{r}$ by $\overline{r}\epsilon$, and this process is reversible. $\square$

Since $\mathrm{ad}\,\overline{r} = \mathrm{ad}^0\overline{r} \oplus \mathbf{1}$, this decomposes as

$$H^1(G_{F_v}, \mathrm{ad}\,\overline{r}) \cong H^1(G_{F_v}; \mathbf{F}) \oplus H^1(G_{F_v}; \mathrm{ad}^0\overline{r}).$$

The Fontaine-Laffaille conditions says that

$$\overline{r}|_{G_{F_v}} = \mathbf{G}(M)$$

for M in $\mathcal{MF}/\mathbf{F}$, and $\mathbf{G}$ the Fontaine-Laffaille functor (§5.3). We consider $\mathrm{Ext}^1_{\mathcal{MF}/\mathbf{F}}(M,M)$, which means the group of extensions up to equivalence (we aren't ready to make a statement that there are "enough injectives" in $\mathcal{MF}/\mathbf{F}$, which would be necessary to define derived functors). The functor $\mathbf{G}$ induces a map

$$\mathrm{Ext}^1_{\mathcal{MF}/\mathbf{F}}(M,M) \to \mathrm{Ext}^1_{\mathbf{F}[G_{F_v}]}(\overline{r}, \overline{r}). \tag{7.2.3}$$

The Fontaine-Laffaille condition means that the extension is in the image of this map. The image of (7.2.3) is denoted $H^1_f(G_{F_v}, \mathrm{ad}^0\overline{r})$ under the identification (7.2.2).

It is a fact, which is maybe not adequately treated in the literature[9], that

$$H^1_f(G_{F_v}, \mathrm{ad}\,\overline{r}) \cong H^1_f(G_{F_v}, \mathrm{ad}^0\overline{r}) \oplus \underbrace{H^1(G_{k(v)}, \mathbf{F})}_{1-\mathrm{dim}}$$

and $H^1_f(G_{F_v}, \mathrm{ad}^0\overline{r}) = H^1(G_{F_v}, \mathrm{ad}^0\overline{r}) \cap H^1_f(G_{F_v}, \mathrm{ad}\,\overline{r})$. Hence

$$\dim_{\mathbf{F}} H^1_f(G_{F_v}, \mathrm{ad}^0 r) = \dim_{\mathbf{F}} \mathrm{Ext}^1_{\mathcal{MF}/\mathbf{F}}(M,M) - 1.$$

Therefore, the condition that $[\phi] \in H^1(G_{F,Q\cup\{v\}}, \mathrm{ad}^0\overline{r})$ is Fontaine-Laffaille is the condition that $[\phi]$ lies in

$$\ker\left(H^1(G_{F,Q\cup\{v\}}, \mathrm{ad}^0\overline{r}) \to \bigoplus_{v\mid\ell} H^1(G_{F_v}, \mathrm{ad}^0\overline{r})/H^1_f(G_{F_v}, \mathrm{ad}^0\overline{r})\right).$$

7.3. **Dimension calculations.** We are going to work towards calculating the dimension of (7.2.1).

[9]The problem here is that the tensor product of two FL modules needn't be FL again (the natural tensor product filtration can become too long), but when it is it should be the case that $\mathbf{G}(M\otimes M') = \mathbf{G}(M)\otimes\mathbf{G}(M')$.

7.3.1. *The case $v \mid \ell$: Fontaine-Laffaille theory.* We now want to compute the dimension of the Galois cohomology groups $H^1_f(G_{F_v}, \mathrm{ad}^0 \overline{r})$.

Let's explicate the definition of the Fontaine-Laffaille category $\mathcal{MF}/\mathbf{F}$. When working with $\mathbf{F}$-coefficients, we find that the definition in §5.3 amounts to just:

- M an $\mathbf{F} \otimes_{\mathbf{Z}_\ell} \mathcal{O}_{F_v}$-module, with
- a filtration Fil^i such that $\mathrm{Fil}^0 = M$ and $\mathrm{Fil}^{\ell-1} = 0$, and
- a map
$$\Phi^\bullet \colon \mathrm{gr}^\bullet M \xrightarrow{\sim} M$$
which is $\mathrm{Id} \otimes \mathrm{Frob}_\ell^{-1}$-linear.

Given an extension $E \in \mathrm{Ext}^1_{\mathcal{MF}/\mathbf{F}}(M, M)$
$$0 \to M \to E \to M \to 0,$$
we can pick a splitting $s \colon M \to E$ which preserve the filtration. That is, s induces $E \cong M \oplus M$ such that $\mathrm{Fil}(E) = \mathrm{Fil}(M) \oplus \mathrm{Fil}(M)$. To complete this to an FL structure, we need to give a map $\Phi_E \in \mathrm{Hom}^{\mathrm{Frob}}_{\mathbf{F} \otimes \mathcal{O}_{F_v}}(\mathrm{gr}^\bullet E, E)$ which is $\mathrm{Id} \otimes \mathrm{Frob}_\ell^{-1}$-linear. It is required to have the form
$$\Phi_E = \begin{pmatrix} \Phi_M & * \\ & \Phi_M \end{pmatrix}$$
so it is specified by the upper-right hand corner, which is a $\Phi \in \mathrm{Hom}^{\mathrm{Frob}}_{\mathbf{F} \otimes \mathcal{O}_{F_v}}(\mathrm{gr}^\bullet M, M)$. Thus, we have a surjection
$$\mathrm{Hom}^{\mathrm{Frob}}_{\mathbf{F} \otimes \mathcal{O}_{F_v}}(\mathrm{gr}^\bullet M, M) \to \mathrm{Ext}^1_{\mathcal{MF}/\mathbf{F}}(M, M) \to 0.$$

The different splittings which induce the same E are a torsor for homomorphisms $M \to M$ that preserve the filtration, i.e. $\mathrm{Fil}^0 \mathrm{Hom}(M, M)$. So extend to an exact sequence.
$$\mathrm{Fil}^0 \mathrm{Hom}(M, M) \to \mathrm{Hom}^{\mathrm{Frob}}_{\mathbf{F} \otimes \mathcal{O}_{F_v}}(\mathrm{gr}^* M, M) \to \mathrm{Ext}^1_{\mathcal{MF}/\mathbf{F}}(M, M) \to 0.$$
A change in the choice of splitting amounts to multiplying E by a matrix of the form
$$\begin{pmatrix} 1 & \psi \\ & 1 \end{pmatrix}$$
which conjugates Φ_E:
$$\begin{pmatrix} 1 & \psi \\ & 1 \end{pmatrix} \begin{pmatrix} \Phi_M & \Phi \\ & \Phi_M \end{pmatrix} \begin{pmatrix} 1 & -\psi \\ & 1 \end{pmatrix} = \begin{pmatrix} \Phi_M & \Phi + \psi \Phi_M - \Phi_M \psi \\ & \Phi_M \end{pmatrix}.$$
Therefore, the kernel of $\mathrm{Fil}^0 \mathrm{Hom}(M, M) \to \mathrm{Hom}^{\mathrm{Frob}}_{\mathbf{F} \otimes \mathcal{O}_{F_v}}(\mathrm{gr}^* M, M)$ is the space of $\psi \colon M \to M$ that commute with Φ_M. This is just the space $\mathrm{Hom}_{\mathcal{MF}/\mathbf{F}}(M, M)$. So we've produced an exact sequence
$$0 \to \mathrm{Hom}_{\mathcal{MF}/\mathbf{F}}(M, M) \to \mathrm{Fil}^0 \mathrm{Hom}_{\mathbf{F} \otimes \mathcal{O}_{F_v}}(M, M) \to \mathrm{Hom}_{\mathbf{F} \otimes \mathcal{O}_{F_v}}(\mathrm{gr}^\bullet M, M) \to \mathrm{Ext}^1_{\mathcal{MF}/\mathbf{F}}(M, M) \to 0.$$

What are the dimensions?

- We have $\dim_{\mathbf{F}} \mathrm{Hom}^{\mathrm{Frob}}_{\mathbf{F} \otimes \mathcal{O}_{F_v}}(\mathrm{gr}^* M, M) = 4[k(v) : \mathbf{F}]$. This is because, after choosing a basis for M, such a homomorphism can be specified by a 2×2 matrix with entries in $k(v)$, but then we take the dimension over $\mathbf{F}$.

- The assumption of distinct Hodge-Tate weights (and the fact that $\dim_{\mathbf{F}} \overline{r} = 2$) implies that M has a filtration with two steps, so $\mathrm{Fil}^0 \operatorname{Hom}(M, M)$ is the space of upper-triangular matrices with entries in $k(v)$, hence has dimension $3[k(v) : \mathbf{F}]$.
- Finally, using the equivalence $\operatorname{Hom}_{\mathscr{MF}/\mathbf{F}}(M, M) \cong \operatorname{Hom}_{G_{F_v}}(\overline{r}, \overline{r})$ we find that it has $\mathbf{F}$-dimension $1 + \dim H^0(G_{F_v}, \mathrm{ad}^0 \overline{r})$.

In conclusion (using that the Euler characteristic of an exact sequence is 0), we've found that

$$\dim H^1_f(G_{F_v}\ \mathrm{ad}^0 \overline{r}) = [F_v : \mathbf{Q}_\ell] + \dim H^0(G_{F_v}, \mathrm{ad}^0 \overline{r}). \tag{7.3.1}$$

7.3.2. *Selmer group formalism.* Let T be a finite set of primes of a number field F. Let M be an $\mathscr{O}[G_{F,T}]$-module which is finite over $\mathscr{O}$. Suppose that if $p \mid \#M$ then all $v \mid p$ are in T.

For all $v \mid T$, we suppose given a subspace $L_v \subset H^1(G_{F_v}, M)$. Set $\mathscr{L} = \{L_v\}$.

Definition 7.3.1. We define $H^1_{\mathscr{L}}(G_{F,T}; M) \subset H^1(G_{F,T}; M)$ to be the subspace of elements whose image under the local restriction maps lie in L_v, i.e.

$$H^1_{\mathscr{L}}(G_{F,T}; M) = \ker\left(H^1(G_{F,T}; M) \to \bigoplus_{v \in T} H^1(G_{F_v}, M)/L_v \right).$$

For $M = \mathrm{ad}^0 \overline{r}$, we want to take

$$\mathscr{L}_v = \begin{cases} H^1_f(G_{F_v}, \overline{r}) & v \mid \ell, \\ H^1(G_{F_v}, \mathrm{ad}^0 \overline{r}) & v \in Q. \end{cases}$$

Upshot: by §7.1, R_Q^{univ} is topologically generated by $\dim H^1_{\mathscr{L}}(G_{F,Q\cup\{\ell\}}, \mathrm{ad}^0 \overline{r})$ elements.

Consider what happens if we change T to $T \cup \{v\}$, and we take

$$L_v = H^1(G_{k(v)}; M) \subset H^1(G_{F_v}; M).$$

Then we would have

$$H^1_{\mathscr{L}}(G_{F,T}; M) \to H^1_{\mathscr{L}\cup\{L_v\}}(G_{F,T\cup\{v\}}; M)$$

since the condition L_v imposes unramifiedness.

7.3.3. *Galois cohomology of local fields.* The cohomology of a local field behaves like the cohomology of a compact 2-manifold. For example, if v is a finite place then we have

$$H^i(G_{F_v}; M) = 0, \quad i > 2$$
$$H^2(G_{F_v}; \mathbf{F}(\epsilon_\ell)) \cong \mathbf{F}$$

and there is a form of Poincaré duality, which says that for $M^*(1) := \operatorname{Hom}_{\mathbf{F}}(M, \mathbf{F}(\epsilon_\ell))$ the pairing

$$H^i(G_{F_v}; M) \times H^{2-i}(G_{F_v}, M^*(1)) \to H^2(G_{F_v}; \mathbf{F}(\epsilon_\ell)) \xrightarrow{\sim} \mathbf{F}$$

is perfect.

We regard H^0 as being "easy". Then Poincaré duality makes H^2 similarly "easy". So the only mysterious group is H^1. However, information can be gotten from the Euler characteristic formula.

Theorem 7.3.2 (Tate). *Let M be a $\mathbf{F}$-vector space. We have*

$$\dim_{\mathbf{F}} H^1(G_{F_v};M) = \dim_{\mathbf{F}} H^0(G_{F_v};M) + \dim_{\mathbf{F}} H^0(G_{F_v};M^*(1)) + \begin{cases} (\dim_{\mathbf{F}} M)[F_v:\mathbf{Q}_\ell] & v \mid \ell \\ 0 & v \nmid \ell \end{cases}.$$

Given $L_v \subset H^i(G_{F_v};M)$, we can define $L_v^\perp \subset H^{2-i}(G_{F_v}, M^*(1))$ to be its annihilator. A key property of this construction is that

$$H^1(G_{k(v)};M)^\perp = H^1(G_{k(v)};M^*(1)).$$

Given $\mathscr{L}_v = (L_v)$, we then define $\mathscr{L}^\perp := (L_v^\perp)$.

7.3.4. *Galois cohomology of global fields.* The cohomology of the groups $G_{F,T}$ behaves like that of a 3-manifold with boundary, having boundary components for each $v \in T$ and each $v \mid \infty$. (For $v \in T$, the boundary components have cohomology like that of G_{F_v}.)

Poincaré duality for manifolds with boundary is more complicated, but whatever the statements are, analogous ones are true here. In particular, Tate proved global duality and Euler characteristic formulae. We'll just state what we need (combining the duality and Euler characteristic formulae).

Theorem 7.3.3 (Poitou-Tate). *We have*

$$\begin{aligned} -\dim H^0(G_{F,T};) + \dim H^1_{\mathscr{L}}(G_{F,T};M) - \dim H^1_{\mathscr{L}^\perp}(G_{F,T};M^*(1)) + \dim H^0(G_{F,T};M^*(1)) \\ = -\sum_{v\mid\infty} \dim M^{G_{F_v}} + \sum_{v\in T}(\dim L_v - \dim H^0(G_{F_v};M)). \end{aligned}$$

It is very difficult to calculate the dimensions of Selmer groups; this gives some information.

7.3.5. *Putting things together.* We have seen that R_Q^{univ} is topologically generated by $\dim_{\mathbf{F}} H^1_{\mathscr{L}}(G_{F,Q\cup\{\ell\}};\mathrm{ad}^0\,\overline{r})$ elements, where $\mathscr{L}$ has

$$L_v = \begin{cases} H^1_f(G_{F_v};\mathrm{ad}^0\,\overline{r}) & v \mid \ell, \\ H^1(G_{F_v};\mathrm{ad}^0\,\overline{r}) & v \in Q. \end{cases}$$

- At $v \mid \infty$, we have $\overline{r}(c) \sim \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$ hence

$$(\mathrm{ad}^0\,\overline{r})(c) \sim \begin{pmatrix} 1 & & \\ & -1 & \\ & & -1 \end{pmatrix}$$

  so this contributes 1 for each such v.
- For $v \mid \ell$, we found $[F_v:\mathbf{Q}_\ell] + \dim H^0(G_{F_v};\mathrm{ad}^0\,\overline{r})$.
- Finally, for $v \in Q$ we get $\dim H^1(G_{F_v};\mathrm{ad}^0\,\overline{r}) - \dim H^0(G_{F_v};\mathrm{ad}^0\,\overline{r})$. Since $v \nmid \ell$, we apply Theorem 7.3.2 to get that this is

$$\dim H^2(G_{F_v};\mathrm{ad}^0\,\overline{r}) = \dim H^0(G_{F_v};(\mathrm{ad}^0\,\overline{r})^*(1)).$$

The assumption $v \in Q$ implies $\overline{\epsilon}_\ell(\mathrm{Frob}_v) = 1$, so we can ignore the Tate twist. Since

$$(\mathrm{ad}^0\,\overline{r})(\mathrm{Frob}_v) = \begin{pmatrix} \alpha_v/\beta_v & & \\ & \beta_v/\alpha_v & \\ & & 1 \end{pmatrix}$$

we see a contribution of 1, hence a total contribution of $\#Q$ from such v.

- $\dim H^0(G_{FT}; \mathrm{ad}^0\,\overline{r}) = 0$, because $\overline{r}$ is irreducible.
- $\dim H^0(G_{FT}; (\mathrm{ad}^0\,\overline{r})(1)) = 0$, because we can ignore the twist by passing up to $G_{F(\zeta_\ell)}$, and $\overline{r}|_{G_{F(\zeta_\ell)}}$ is irreducible.

The upshot is that

$$\dim_{\mathbf{F}} H^1_{\mathscr{L}}(G_{F,Q\cup\{\ell\}}; \mathrm{ad}^0\,\overline{r}) = \dim_{\mathbf{F}} H^1_{\mathscr{L}^\perp}(G_{F,Q\cup\{\ell\}}; (\mathrm{ad}^0\,\overline{r})(1)) + \#Q. \tag{7.3.2}$$

7.4. **Preparations for patching.**

7.4.1. *Philosophy.* For $Q = \emptyset$, if $h^1 = \dim H^1_{\mathscr{L}_\emptyset}(G_{F,\{\ell\}}; \mathrm{ad}^0\,\overline{r})$ and $h^2 = \dim H^1_{\mathscr{L}_\emptyset^\perp}(G_{F,\{\ell\}}; \mathrm{ad}^0\,\overline{r}(1))$, then (7.3.2) gives $h^1 = h^2$ and §7.1 shows that the universal deformation ring has the form

$$R^{\mathrm{univ}}_\emptyset = \mathscr{O}[[x_1, \dots, x_h]]/(f_1, \dots, f_{h^2}).$$

Then you get that $R^{\mathrm{univ}} = \mathscr{O}$, and you have $R^{\mathrm{univ}} = \mathscr{O} \twoheadrightarrow \mathbf{T}$ must be an isomorphism since $\mathbf{T}$ is free over $\mathscr{O}$.

We're going to choose Q so that h^1 and h^2 don't change, but $f_1, \dots, f_{h^2}$ become "deeper" in the augmentation filtration. The idea is that in a "limiting situation" (whatever that means; a priori these rings for different Q don't map to each other), they would become 0, and R^{univ}_∞ would become $\mathscr{O}[[x_1, \dots, x_{h^1}]]$. This would map to $\mathbf{T}_\infty$, and it would automatically be an isomorphism if you knew that $\dim \mathbf{T}_\infty = h^1 + 1$. This largeness of dimension will come from the largeness of the $\mathbf{T}_Q$-module S_Q.

7.4.2. *Selection of Taylor-Wiles primes.* The strategy we just explained requires that (7.3.2) stay constant. So we need to choose Q so that $\dim H^1_{\mathscr{L}_Q^\perp}$ drops as we make $\#Q$ bigger.

Proposition 7.4.1. *Let* $s = \dim H^1_{\mathscr{L}_\emptyset^\perp}(G_{F,\{\ell\}}; \mathrm{ad}^0(\overline{r})(1))$. *For all* $m > 0$ *there exists* Q_m *a set of primes of* F *such that*

1. *For* $v \in Q_m$, *we have* $\#k(v) \equiv 1 \pmod{\ell^m}$, *(so as* m *grows, the pro-*ℓ *quotient* $\Delta_v \twoheadrightarrow \mathbf{Z}/\ell^m\mathbf{Z}$ *gets bigger)*
2. $\overline{r}(\mathrm{Frob}_v)$ *has distinct eigenvalues* $\alpha_v \neq \beta_v$
3. $\#Q_m = s$,
4. R^{univ}_Q *is topologically generated by* r *elements over* $\mathscr{O}$ *(note that* $R^\perp_\emptyset$ *is generated by this same number of elements).*

Remark 7.4.2. Under (1)-(3), (4) is equivalent to

$$\dim H^1_{\mathscr{L}^\perp_{Q_m}}(G_{F,Q_m\cup\{\ell\}}, \mathrm{ad}^0\,\overline{r})(1)) = 0.$$

Proof. We have to understand the cohomology group $H^1_{\mathscr{L}^\perp_{Q_m}}(G_{F,Q_m\cup\{\ell\}}, \mathrm{ad}^0\,\overline{r})(1))$. Let's begin by explicating $\mathscr{L}^\perp_{Q_m}$. For $v\in Q_m$, we have $\mathscr{L}^\perp_{Q_m,v}=0$.

By definition, we have an exact sequence

$$0\to H^1_{\mathscr{L}^\perp_{Q_m}}(G_{F,Q_m\cup\{\ell\}};\mathrm{ad}^0\,\overline{r}(1))\to H^1_{\widetilde{\mathscr{L}}^\perp}(G_{F,Q_m\cup\{\ell\}};\mathrm{ad}^0\,\overline{r}(1))\to\bigoplus_{v\in Q_m}H^1(G_{k(v)};\mathrm{ad}^0(\overline{r})(1)) \tag{7.4.1}$$

where

$$\widetilde{\mathscr{L}_v}=\begin{cases}H^1_f(G_{F_v};\mathrm{ad}^0\,\overline{r}) & v\mid\ell,\\ H^1(G_{k(v)};\mathrm{ad}^0\,\overline{r}) & v\in Q_m,\end{cases}$$

hence

$$H^1_{\widetilde{\mathscr{L}}^\perp}(G_{F,Q_m\cup\{\ell\}};\mathrm{ad}^0\,\overline{r}(1))=H^1_{\mathscr{L}^\perp_\emptyset}(G_{F,\{\ell\}};(\mathrm{ad}^0\,\overline{r})(1)).$$

So we can rewrite (7.4.1) as

$$0\to H^1_{\mathscr{L}^\perp_{Q_m}}(G_{F,Q_m\cup\{\ell\}};(\mathrm{ad}^0\,\overline{r})(1))\to H^1_{\mathscr{L}^\perp_\emptyset}(G_{F,\{\ell\}};(\mathrm{ad}^0\,\overline{r})(1))\to\bigoplus_{v\in Q_m}H^1(G_{k(v)};(\mathrm{ad}^0\,\overline{r})(1)).$$

Since $G_{k(v)}\cong\widehat{\mathbf{Z}}$, we can understand its Galois cohomology easily:

$$H^1(G_{k(v)};\mathrm{ad}^0(\overline{r}))\cong\frac{(\mathrm{ad}^0\,\overline{r})(1)}{(\mathrm{Frob}_v-1)(\mathrm{ad}^0\,\overline{r})(1)}.$$

Since by assumption Frob_v acts on $\mathrm{ad}^0(\overline{r})$ with eigenvalues $\alpha_v/\beta_v,\beta_v/\alpha_v,1$ with $\alpha_v\neq\beta_v$, we have $\dim_{\mathbf{F}}H^1(G_{k(v)};\mathrm{ad}^0(\overline{r}))=1$. The assumption $\#k(v)\equiv 1 \pmod{\ell^m}$ lets us ignore the Tate twist; we'll use this throughout to simplify the notation.

7.4.3. It's easy to pick a bunch of v satisfying (2) and (3). To satisfy (4), we will find Q_m so that

$$H^1_{\mathscr{L}^\perp_\emptyset}(G_{F,\{\ell\}};(\mathrm{ad}^0\,\overline{r}))\hookrightarrow\bigoplus_{v\in Q_m}H^1(G_{k(v)};\mathrm{ad}^0(\overline{r})). \tag{7.4.2}$$

We can then throw away extra elements of Q_m to achieve (1) as well.

To prove (7.4.2), we want to show that for all non-zero $\phi\in H^1_{\mathscr{L}^\perp_\emptyset}(G_{F,\{\ell\}};(\mathrm{ad}^0\,\overline{r}))$, we can find v such that ϕ restricts to a non-zero class in $H^1(G_{k(v)};\mathrm{ad}^0(\overline{r}))$, i.e. $\phi(\mathrm{Frob}_v)\notin(\mathrm{Frob}_v-1)\,\mathrm{ad}^0\,\overline{r}$. We'll do this using Chebotarev's density theorem.

7.4.4. So it suffices to know that there exists *some* $\sigma\in G_{F(\zeta_{\ell^m})}$ such that

(i) $\overline{r}(\sigma)\neq 1$ and has order coprime to ℓ (to get distinct eigenvalues),
(ii) $\phi(\sigma)\notin(\sigma-1)\,\mathrm{ad}^0\,\overline{r}$.

We'll now digest this.

7.4.5. Write $D=\overline{F}^{\ker\mathrm{ad}\,\overline{r}}$, a finite extension of F that trivializes $\mathrm{ad}\,\overline{r}$. Well argue that it suffices to produce $\sigma\in\mathrm{Gal}(D(\zeta_{\ell^m})/F(\zeta_{\ell^m}))$ such that

(i) $\sigma\neq 1$ and has order coprime to ℓ,
(ii) $\mathrm{ad}^0\,\sigma$ has eigenvalue 1 on $\langle\phi(G_{D(\zeta_{\ell^m})})\rangle\subset\mathrm{ad}^0\,\overline{r}$.

Why? Pick a lift $\widetilde{\sigma} \in G_{F(\zeta_{\ell m})}$ of σ. Then take any $\tau \in G_{D(\zeta_{\ell m})}$ and consider

$$\phi(\tau\widetilde{\sigma}) = \phi(\tau) + \phi(\widetilde{\sigma}).$$

(We have $\tau \cdot \phi(\widetilde{\sigma}) = \phi(\widetilde{\sigma})$ by the definition of D.) If $\phi(\widetilde{\sigma}) \notin (\sigma - 1)\mathrm{ad}^0\,\overline{r}$, then we win. Otherwise $\phi(\widetilde{\sigma}) \in (\sigma - 1)\mathrm{ad}^0\,\overline{r}$, and we want to find τ such that $\phi(\tau) \notin (\sigma-1)\mathrm{ad}^0\,\overline{r}$, for then we can put $\tau\widetilde{\sigma}$ into §7.4.4. So if $\sigma - 1$ is not invertible on $\mathrm{ad}^0\,\overline{r}$, then we win.

7.4.6. Let's show that we can achieve (ii) in §7.4.5. What's the content here? We're in trouble if $\langle \phi(G_{D(\zeta_{\ell m})}) \rangle = 0$. We'll show that the restriction map

$$H^1_{\mathscr{L}_\emptyset^\perp}(G_{F,\{\ell\}}; (\mathrm{ad}^0\,\overline{r})(1)) \hookrightarrow H^1(G_{D(\zeta_{\ell m})}, (\mathrm{ad}^0\,\overline{r})(1)) \tag{7.4.3}$$

which rules this out. The kernel comes from $H^1(\mathrm{Gal}(D(\zeta_{\ell m})/F), \mathrm{ad}^0\,\overline{r}(1)) = 0$, so we'll arrange that this is 0.

Then we'll show that for any non-zero $\mathrm{Gal}(D(\zeta_{\ell m})/F)$-submodule $W \subset \mathrm{ad}^0\,\overline{r}$, there exists $\sigma \in \mathrm{Gal}(D(\zeta_{\ell m})/F(\zeta_{\ell m}))$ such that

- $\sigma \neq 1$ has order prime to ℓ, and
- σ has an eigenvalue 1 on W.

7.4.7. We prove the second statement first. We know that $\mathrm{ad}^0\,\overline{r}|_{G_{F(\zeta_\ell)}}$ is semi-simple, by a criterion of Serre (which says that a tensor product of semisimple representations is semisimple if the dimension is small relative to ℓ), because $2-1+2-1 < \ell$.

Then we have the the following possibilities:

- $\mathrm{ad}^0\,\overline{r}|_{G_{F(\zeta_\ell)}}$ is irreducible. In this case, any $\sigma \in \mathrm{Gal}(D(\zeta_\ell)/F(\zeta_\ell))$ of order $\neq 1$ and coprime to ℓ will do. (If $\mathrm{Gal}(D(\zeta_\ell)/F(\zeta_\ell))$ is an ℓ-group, then $(\mathrm{ad}^0\,\overline{r})|_{G_{F(\zeta_\ell)}}$ is reducible.)
- Otherwise, $\mathrm{ad}^0\,\overline{r}|_{G_{F(\zeta_\ell)}}$ contains a 1-dimensional irreducible character ψ, which forces $\overline{r} \cong \overline{r} \otimes \psi$ over $G_{F(\zeta_\ell)}$, with $\psi^2 = 1$ (by looking at determinants) and $\psi \neq 1$ (because $\overline{r}|_{G_{F(\zeta_\ell)}}$ is irreducible, and we've already stripped out the scalar automorphisms in passing from $\mathrm{ad}\,\overline{r}$ to $\mathrm{ad}^0\,\overline{r}$).

  Then there is a quadratic extension $K/F(\zeta_\ell)$ (cut out by ψ) such that $\overline{r} \cong \mathrm{Ind}_{G_K}^{G_{F(\zeta_\ell)}} \theta$, and satisfying $\theta/\theta^\tau \neq 1$ where $\tau \in \mathrm{Gal}(K/F(\zeta_\ell))$ is the nontrivial element. Then

  $$\mathrm{ad}^0\,\overline{r} = \psi \oplus \mathrm{Ind}_{G_K}^{G_{F(\zeta_\ell)}}(\theta/\theta^\tau).$$

  - For $W = \mathbf{F}(\psi)$, any $\sigma \in G_K$ with $\theta(\sigma) \neq 1$ will do.
  - If $W = \mathrm{Ind}_{G_K}^{G_{F(\zeta_\ell)}} \theta/\theta^\tau$, any $\sigma \in G_{F(\zeta_\ell)} - G_K$ acts through $\overline{r}$ by a matrix of the form

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

    hence has eigenvalues $\pm\alpha$. Then in $\mathrm{ad}^0\,\overline{r}$ it has eigenvalues $-1, -1, +1$ and $\psi(\sigma) = -1$, so there is a $+1$-eigenvalue in W.
  - Finally, $\mathrm{Ind}_{G_K}^{G_{F(\zeta_\ell)}} \theta/\theta^\tau$ could be reducible and W be 1-dimensional again. Then $W \cong \mathbf{F}(\psi')$ and you argue as above.

7.4.8. Now we show that $H^1(\mathrm{Gal}(D(\zeta_{\ell^m})/F);\mathrm{ad}^0\overline{r}(1))=0$, as promised in §7.4.6. We use the inflation-restriction sequence repeatedly. First we look at field extensions

$$F\hookrightarrow D(\zeta_\ell)\hookrightarrow D(\zeta_{\ell^m}).$$

This gives the inflation-restriction sequence

$$0\to H^1(\mathrm{Gal}(D(\zeta_\ell)/F),\mathrm{ad}^0\overline{r}(1))\to H^1(\mathrm{Gal}(D(\zeta_{\ell^m})/F),\mathrm{ad}^0\overline{r}(1))\to H^1(\mathrm{Gal}(D(\zeta_{\ell^m})/D(\zeta_\ell));\mathrm{ad}^0\overline{r}(1))^{G_F}.$$

We first analyze the last term. Since $\mathrm{Gal}(D(\zeta_{\ell^m})/D(\zeta_\ell))$ acts trivially on $\mathrm{ad}^0\overline{r}(1)$, it can be understood as

$$\mathrm{Hom}_{G_F}(\mathrm{Gal}(D(\zeta_{\ell^m})/D(\zeta_\ell));\mathrm{ad}^0\overline{r}(1))$$

Now, G_F acts trivially on $\mathrm{Gal}(D(\zeta_{\ell^m})/D(\zeta_\ell))$, since the latter injects into $\mathrm{Gal}(F(\zeta_{\ell^m}/F(\zeta_\ell))$, and $F(\zeta_{\ell^m})$ is abelian over F. So this is just the same as

$$\mathrm{Hom}(\mathrm{Gal}(D(\zeta_{\ell^m})/D(\zeta_\ell));\mathrm{ad}^0\overline{r}(1)^{G_F}).$$

Next, note that $\mathrm{ad}^0\overline{r}(1)^{G_F}=0$ otherwise we would have an isomorphism $\overline{r}\cong\overline{r}\otimes\epsilon_\ell$ with trace 0, but we assumed $\overline{r}|_{G_{F(\zeta_\ell)}}$ is irreducible.

It remains to show that $H^1(\mathrm{Gal}(D(\zeta_\ell)/F),\mathrm{ad}^0\overline{r}(1))=0$. For this we consider the inflation-restriction sequence

$$0\to H^1(\mathrm{Gal}(D/F),\mathrm{ad}^0\overline{r}(1)^{G_D})\to H^1(\mathrm{Gal}(D(\zeta_\ell/F),\mathrm{ad}^0\overline{r}(1))\to H^1(\mathrm{Gal}(D(\zeta_\ell)/D),\mathrm{ad}^0\overline{r}(1)).$$

The last term $H^1(\mathrm{Gal}(D(\zeta_\ell)/D),\mathrm{ad}^0\overline{r}(1))$ is 0 because $\mathrm{Gal}(D(\zeta_\ell)/D)$ has order prime to ℓ. For the first term, G_D acts only through the cyclotomic character so we have $\mathrm{ad}^0\overline{r}(1)^{G_D}=00$ unless $\zeta_\ell\in D$, in which case the first term becomes $H^1(\mathrm{Gal}(D/F),\mathrm{ad}^0\overline{r}(1))$.

We know something about $\mathrm{Gal}(D/F)$, namely that it embeds in $\mathrm{PGL}_2(\overline{\mathbf{F}}_\ell)$. The possibilities for its image are

$$A_4\quad S_4\quad A_5\quad \mathrm{PSL}_2(\mathbf{F}_{\ell^r})\quad \mathrm{PGL}_2(\mathbf{F}_{\ell^r})\quad D_{2s}\ (\text{if }\ell\nmid s)$$

Furthermore, if $\zeta_\ell\in D$ then $\mathrm{Gal}(D/F)$ surjects onto $\mathrm{Gal}(F(\zeta_\ell)/F)$. We examin the maximal abelian quotients of the above groups:

A_4	S_4	A_5	$\mathrm{PSL}_2(\mathbf{F}_{\ell^r})$	$\mathrm{PGL}_2(\mathbf{F}_{\ell^r})$	D_{2s} (if $\ell\nmid s$)
C_3	C_2	$\{1\}$	$\{1\}$ (unless $(\ell,r)=(3,1)$)	$\{\pm1\}$	$\{\pm1\}$ or $\{\pm1\}^2$

Since F is totally real, the Galois group in question is cyclic of even degree. So in the case $\zeta_\ell\in D$ we must have $\mathrm{Gal}(D/F)\cong S_4$ or $\mathrm{PGL}_2(\mathbf{F}_{\ell^r})$ or D_{2s}, and $\mathrm{Gal}(F(\zeta_\ell)/F)\cong C_2$. If $\ell>3$ is unramified in F, then $[F(\zeta_\ell):F]=\ell-1$, a contradiction.

If $\ell=3$ (so $S_4=\mathrm{PGL}_2(\mathbf{F}_\ell)$, which we include because of its historical significance, then we get $H^1(\mathrm{Gal}(D/F);\mathrm{ad}^0\overline{r}(1))=0$ if $\mathrm{Gal}(D/F)=D_{2s}$. Otherwise, we use the group-theoretic fact that $H^1(\mathrm{PGL}_2(\mathbf{F}_{\ell^r}),\mathrm{ad}^0\overline{r}(1))=0$ unless $\ell^r=5$.

□

8. PATCHING

8.1. Summary. Let's summarize where we are.

Write $R_\infty = \mathcal{O}[[X_1, \dots, X_s]]$ and $\Lambda = \mathcal{O}[[\mathbf{Z}_\ell^s]] = \mathcal{O}[[S_1, \dots, S_s]]$.

We consider the space of modular forms of level Q_m, denoted S_{Q_m}. This maps down to $S_\emptyset$. It has an action of $\mathcal{O}[\Delta_{Q_m}]$, which admits a surjection from Λ. It also has an action of the Hecke algebra $\mathbf{T}_{Q_m}$. This actions are compatible with the map $R_{Q_m}^{\mathrm{univ}} \twoheadrightarrow \mathbf{T}_{Q_m}$. Since R_{Q_m} is topologically generated by s elements, there exists $R_\infty \twoheadrightarrow R_{Q_m}^{\mathrm{univ}}$.

$$\begin{tikzcd}
 & \Lambda \arrow[r, dashed, two heads] \arrow[dl, dashed] & \mathcal{O}[\Delta_{Q_m}] \arrow[dl] \arrow[d] & & \\
R_\infty \arrow[r, dashed, two heads] & R_{Q_m}^{\mathrm{univ}} \arrow[r, two heads] \arrow[d, two heads] & \mathbf{T}_{Q_m} \arrow[d, two heads] & \curvearrowright & S_{Q_m} \arrow[d, two heads] \\
 & R_\emptyset^{\mathrm{univ}} \arrow[r] & \mathbf{T}_\emptyset & \curvearrowright & S_\emptyset
\end{tikzcd}$$

Here we have used dashed arrows to label "random" ring homorphisms, which exist because the domains are free so we can construct them by making choices, but which have no "natural" interpretation.

Recall that for the augmentation ideal $\mathfrak{a} = \langle \gamma - 1 : \gamma \in \mathbf{Z}_\ell^s \rangle = (S_1, \dots, S_s)$, we know that S_{Q_m} is free over $\mathcal{O}[\Delta_{Q_m}]$ and $S_{Q_m}/\mathfrak{a} \xrightarrow{\sim} S_\emptyset$.

This is already more information than we really need. The key picture to extract is

$$\begin{tikzcd}
\Lambda \arrow[r] & R_\infty \arrow[d, two heads] & & \curvearrowright & S_{Q_m} \\
 & R_\emptyset^{\mathrm{univ}} \arrow[r, two heads] & \mathbf{T}_\emptyset & \curvearrowright & S_\emptyset
\end{tikzcd}$$

and S_{Q_m} is free over $\Lambda / \ker\big(\Lambda \to \mathcal{O}[\Delta_{Q_m}]\big)$. The kernel is contained in $I_m := \langle \gamma - 1 : \gamma \in (\ell^m \mathbf{Z}_\ell)^s \rangle$, which has the property that $I_m \to 0$ in a sense; a precise statement is that

$$\cap_m I_m = 0.$$

So the group rings are better and better approximations to Λ. We want to prove that $R_\emptyset^{\mathrm{univ}} \xrightarrow{\sim} \mathbf{T}_\emptyset$.

8.2. Ultrafilters. We consider the power set $\mathscr{P}(\mathbf{Z}_{\geq 0})$.

Definition 8.2.1. A *filter* is a set $\mathscr{F}$ of subsets of $\mathbf{Z}_{\geq 0}$ such that

(1) $\mathbf{Z}_{\geq 0} \in \mathscr{F}, \emptyset \notin \mathscr{F}$,
(2) If $B \in \mathscr{F}$ and $A \supset B$, then $A \in \mathscr{F}$.
(3) $A, B \in \mathscr{F} \implies A \cap B \in \mathscr{F}$.

An *ultrafilter* has the additional property that

$$\text{if } \mathscr{A} \subset \mathbf{Z}_{\geq 0} \text{ then either: } A \cap \mathscr{F} \text{ or } A^c := \mathbf{Z}_{>0} - \mathscr{A} \in \mathscr{F}. \tag{8.2.1}$$

(Only one of A and A^c can be in $\mathscr{F}$, since their intersection is empty.)

Lemma 8.2.2. *Another way to formulate the ultrafilter property is:*

$$A \cup B \in \mathscr{F} \implies A \in \mathscr{F} \text{ or } B \in \mathscr{F}. \tag{8.2.2}$$

Proof. It is obvious that (8.2.2) implies (8.2.1). Conversely, if $A \cup B \in \mathscr{F}$, we can write

$$A \cup B = (A^c \cap B^c)^c \in \mathscr{F}.$$

Hence $(A^c \cap B^c) \notin \mathscr{F}$, so either $A^c \notin \mathscr{F}$ or $B^c \notin \mathscr{F}$. □

Example 8.2.3. For $m \in \mathbf{Z}_{\geq 0}$, an ultrafilter is $\mathscr{F}_m := \{A \in \mathbf{Z}_{>0} : m \in A\}$. This is a *principal ultrafilter.*

Example 8.2.4. The "cofinite filter" $\mathscr{F}_{\mathrm{cof}} := \{A \subset \mathbf{Z}_{>0} : \#(\mathbf{Z}_{>0} - A) < \infty\}$ is a filter.

Lemma 8.2.5. *An ultrafilter $\mathscr{F}$ is not principal if and only if $\mathscr{F} \supset \mathscr{F}_{\mathrm{cof}}$.*

Proof. The direction $\Longleftarrow$ is obvious. Conversely, if $\mathscr{F}$ doesn't contain the cofinite filter, then there exists $A \in \mathscr{F}$ which is finite. Then A is a finite union of singletons, so by (8.2.2) we get a singleton in $\mathscr{F}$. □

Lemma 8.2.6. *An ultrafilter is a maximal filter.*

Proof. Suppose $\mathscr{F}$ is a maximal filter which is not an ultrafilter. Then there is A such that A and A^c are both not in $\mathscr{F}$. One checks that $\mathscr{F}^+ := \{C \supset A \cap B : B \in \mathscr{F}\}$ is a filter. □

Lemma 8.2.7. *Any filter is contained in an ultrafilter.*

Proof. Zorn's lemma. □

Corollary 8.2.8. *Non-principal ultrafilters exist.*

8.3. **Products of Artinian rings.** The point of ultrafilters is to describe the spectrum of (infinite) products of rings. Suppose for all $m \in \mathbf{Z}_{>0}$ we have a local Artinian ring R_m, with maximal ideal $\mathfrak{m}_m$, and consider $R := \prod_m R_m$.

Given $(x_m) \in R_m$, let

$$Z(x_m) = \{m \in \mathbf{Z}_{>0} : X_m \in \mathfrak{m}_m\}.$$

This construction has the following properties:

(i) $Z((x_m)+(y_m)) \supset Z(x_m) \cap Z(y_m)$ – this boils down to saying that if $x_m, y_m \in \mathfrak{m}_m$ then $x_m + y_m \in \mathfrak{m}_m$.

(ii) $Z((x_m)(y_m)) = Z(x_m) \cup Z(y_m)$ – this boils down to saying that if $x_m y_m \in \mathfrak{m}_m$, then $x_m \in \mathfrak{m}_m$ or $y_m \in \mathfrak{m}_m$.

For a subset $A \subset \mathbf{Z}_{>0}$, we introduce the idempotent

$$(e_A)_m = \begin{cases} 1_m & m \in A, \\ 0_m & m \notin A. \end{cases}$$

This construction has the following properties:

(1) $Z(e_A) = A^c$,
(2) $e_A e_B = e_{A \cap B}$, and
(3) $e_{A \cup B} = e_A + e_B - e_A e_B$.

Given a prime ideal $\mathfrak{p} \subset R$, we define

$$\mathscr{F}(\mathfrak{p}) = \{Z(x_m) : (x_m) \in \mathfrak{p}\}.$$

Lemma 8.3.1. *We have* $\mathscr{F}(\mathfrak{p}) = \{A^c : e_A \in \mathfrak{p}\}$.

Proof. The containment $\supset$ is obvious. The converse is also quite obvious – from the perspective of applying Z, we may as well assume replace any element of R by one with only 0_m's and 1_m's. □

Lemma 8.3.2. *For a prime ideal* $\mathfrak{p} \subset R$, $\mathscr{F}(\mathfrak{p})$ *is an ultrafilter.*

Proof. We check off the list of conditions in Definition 8.2.1. For (1), we use that $0 \in \mathfrak{p}$ and $1 \notin \mathfrak{p}$. For (2), we use that $e_B \in \mathfrak{p} \implies e_A \in \mathfrak{p}$ for $A \subset B$, hence $B^c \in \mathscr{F}(\mathfrak{p}) \implies A^c \in \mathscr{F}(\mathfrak{p})$. For (3), if $e_A \in \mathfrak{p}, e_B \in \mathfrak{p}$ then $e_{A \cup B} \in \mathfrak{p}$, so $A^c \cap B^c \in \mathscr{F}(\mathfrak{p})$.

For the ultrafilter property, we observe that

$$A \cup B \in \mathscr{F}(\mathfrak{p}) \implies e_{A^c \cap B^c} = e_{A^c} e_{B^c} \in \mathfrak{p} \implies e_{A^c} \in \mathfrak{p} \text{ or } e_{B^c} \in \mathfrak{p}.$$

□

Conversely, if $\mathscr{F}$ is an ultrafilter then we define

$$\mathfrak{p}(\mathscr{F}) = \{(x_m) \in R : Z(x_m) \in \mathscr{F}\}.$$

One checks that this is a prime ideal.

Theorem 8.3.3. *This induces a bijection between ultrafilters on* $\mathbf{Z}_{\geq 0}$ *and prime ideals of* R.

We'll be interested in this when R_m is independent of m, and has residue field $\mathbf{F}$.

Example 8.3.4. The point is that there are some surprising points of R. For example, for $R = \prod_p \mathbf{F}_p$ the principal ultrafilters correspond to projection to the factor p. However, we know that there are non-principal ultrafilters, so let $\mathfrak{p}$ be the corresponding prime. What does $R/\mathfrak{p}$ look like? We claim that it has characteristic 0.

We have a map $\mathbf{Z} \to R/\mathfrak{p}$, and we want to show that it's injective. If n went to 0, then $(n, n, \ldots, n) \in \mathfrak{p}$. So

$$Z((n, n, \ldots, n)) = \{p : p \mid n\} \in \mathscr{F}(\mathfrak{p}).$$

If $n \neq 0$ then this is a finite set, and a non-principal ultrafilter cannot contain a finite set.

Example 8.3.5. Suppose $R_m = R_0$, independent of m, and $R_0/\mathfrak{m}_0$ is a *finite* field. Then we claim that

$$(\prod R_m)/\mathfrak{p} \cong R_0/\mathfrak{m}_0$$

for any $\mathfrak{p} \in \operatorname{Spec} R$.

We certainly get a map $R_0 \to (\prod R_m)/\mathfrak{p}$, coming from the diagonal map. This obviously sends $\mathfrak{m}_0$ to $\mathfrak{p}$, so factors as

$$R_0/\mathfrak{m}_0 \to (\prod R_m)/\mathfrak{p}.$$

It's automatically injective because $R_0/\mathfrak{m}_0$ is a field. We just need to show surjectivity. Suppose you have $(x_m) \in \prod R_m$. We need to find $\widetilde{y} \in R_0$ such that $(x_m - \widetilde{y}) \in \mathfrak{p}$ for all m. By definition, this is equivalent to

$$Z((x_m - \widetilde{y})) = \{m \colon x_m - \widetilde{y} \in \mathfrak{m}_m\} \overset{?}{\in} \mathscr{F}(\mathfrak{p}).$$

We can rewrite this as

$$\{m \colon x_m \equiv y \pmod{\mathfrak{m}_m}\} \overset{?}{\in} \mathscr{F}(\mathfrak{p}).$$

We can write

$$\mathscr{F} \ni \mathbf{Z}_{>0} = \bigcup_{y \in R_0/\mathfrak{m}_0} \{m \in \mathbf{Z}_{>0} \colon x_m \equiv y \pmod{\mathfrak{m}}\}$$

since every x_m is equivalent to some y.

Since this is just a finite union, as $R_0/\mathfrak{m}_0$ is itself finite, the definition of ultrafilter implies that we can find $y \in R_0/\mathfrak{m}_0$ such that

$$\{m \in \mathbf{Z}_{>0} x_m \equiv y \pmod{\mathfrak{m}}\} \in \mathscr{F}.$$

Example 8.3.6. Suppose $R_m = R_0$ for all m and $\#R_0 < \infty$. Then we claim that for any prime ideal $\mathfrak{p}$ in R, the diagonal map induces an isomorphism

$$R_0 \xrightarrow{\sim} \left(\prod R_m\right)_{\mathfrak{p}}.$$

We need to check that this is injective and surjective.

First we check injectivity. Suppose $x \notin 0$ in R_0 goes to 0, i.e. $(x, x, \ldots)/(1, 1, \ldots) \in \mathfrak{p}$. In other words, there exists $(y_m) \notin \mathfrak{p}$ such that $(y_m x) = 0$. If $y_m \notin \mathfrak{m}_m$ then y_m is a unit, hence $x = 0$. So if $x \neq 0$, then we must have $y_m \in \mathfrak{m}_m$ for all m, hence $Z((y_m)) = \mathbf{Z}_{>0} \in \mathscr{F}$. By definition this implies $(y_m) \in \mathfrak{p}$, contradicting $(y_m) \notin \mathfrak{p}$.

Now for surjectivity. Suppose we're given $(x_m)/(y_m)$ with $Z(y_m) \notin \mathscr{F}$. What does it mean for $x \in R_0$ to map to $(x_m)/(y_m)$? It means that $(x y_m z_m) = (x_m z_m)$ for $(z_m) \notin \mathfrak{p}$, i.e. $Z(z_m) \notin \mathscr{F}$.

For any $x \in R_0$, consider $\{m \colon y_m x = x_m\}$. Since $Z(y_m) \notin \mathscr{F}$ by definition, we have $Z(y_m)^c \in \mathscr{F}$. Now $Z(y_m)^c$ is the set of indices where y_m is a not in the maximal ideal, hence a unit, so for each such index $m \in Z(y_m)^c$ we can invert y_m and find some $x \in R_0$ such that $y_m x = x_m$. This shows that

$$Z(y_m)^c \subset \bigcup_{x \in R_0} \{m \colon y_m x = x_m\}.$$

This is a finite union since R_0 is finite, hence there exists $x \in R_0$ such that $\{m \colon y_m x = x_m\} \in \mathscr{F}$. Fixing such an x. Choose (z_m) such that

$$z_m = \begin{cases} 1 & y_m x = x_m \\ 0 & y_m x \neq x_m \end{cases}$$

which tautologically satisfies $(x y_m z_m) = (x_m z_m)$. Furthermore,

$$Z(z_m) = \mathbf{Z}_0 - \underbrace{\{m \colon y_m x = x_m\}}_{\in \mathscr{F}}$$

so $Z(z_m) \notin \mathscr{F}$.

8.4. **Ultraproduct patching.** We return to the situation

$$\begin{array}{ccccc} \Lambda \longrightarrow R_\infty & & & \curvearrowright & S_{Q_m} \\ \downarrow\!\!\!\!\twoheadrightarrow & & & & \\ R_\emptyset^{\mathrm{univ}} \longrightarrow\!\!\!\!\!\rightarrow \mathbf{T}_\emptyset & & & \curvearrowright & S_\emptyset \end{array}$$

where we set $R_\infty := \mathscr{O}[[X_1,\ldots,X_s]]$ and $\Lambda = \mathscr{O}[[S_1,\ldots,S_s]] \triangleright \mathfrak{a} = (S_1,\ldots,S_s)$. We know that

- $S_{Q_m}/\mathfrak{a} \xrightarrow{\sim} S_\emptyset$ and
- $R_\infty \twoheadrightarrow R_\emptyset^{\mathrm{univ}}$ factors through $R_\infty/\mathfrak{a}$.
- Finally, S_{Q_m} is finite free over $\Lambda/\ker\big(\Lambda \to \mathscr{O}[\Delta_{Q_m}]\big)$.

We want to package these into a statement that looks like the "$m \to \infty$" version of the above, using ultraproducts.

Now, these aren't finite cardinality rings, as we were working with before. But they are inverse limits of rings of finite cardinality. So we choose an open ideal $J \triangleleft \Lambda$; we will mod out by J, which makes everything finite. Then we will take an ultraproduct and localize, and then take an inverse limit over powers of J.

Let $\mathscr{F}$ be a non-principal ultrafilter on $\mathbf{Z}_{>0}$. Consider

$$\prod_m (\mathscr{O}[\Delta_{Q_m}]/J). \tag{8.4.1}$$

Note that $\mathscr{O}[\Delta_{Q_m}]/J \cong \Lambda/J$ for all but finitely many m. We claim that

$$\left(\prod_m (\mathscr{O}[\Delta_{Q_m}]/J)\right)_{\mathfrak{p}(\mathscr{F})} \cong \Lambda/J. \tag{8.4.2}$$

This is almost what we proved before in Example 8.3.6, up to finitely many factors. But changing finitely many factors doesn't affect non-principal ultrafilters – look back at the proof in Example 8.3.6 to see that it still works after changing finitely many factors. The point is that a non-principal ultrafilter behaves likes "$m = \infty$".

Define

$$S_{\infty,J} := \prod_m S_{Q_m}/JS_{Q_m}.$$

This is a module over (8.4.1), so we can localize at $\mathfrak{p}(\mathscr{F})$. By localizing the action of (8.4.1) obtained level-wise, it has an action of (8.4.2):

$$\Lambda/J \curvearrowright S_{\infty,J}.$$

Proposition 8.4.1. *We have the following*

1. $S_{\infty,J}$ *is free over* Λ/J.
2. $S_{\infty,J}/\mathfrak{a} \xrightarrow{\sim} S_\emptyset/JS_\emptyset$.
3. *If* $J \supset J'$ *then the diagram commutes*

$$\begin{array}{ccc} S_{\infty,J'}/J & \xrightarrow{\sim} & S_{\infty,J} \\ \downarrow & & \downarrow \\ S_\emptyset/J & = & S_\emptyset/J \end{array}$$

(4) *R_∞ acts on $S_{\infty,J}$ in a Λ/J-linear manner, and moreover the inflated Λ-action factors through a map $\Lambda \to R_\infty$. For $J \supset J'$, the following diagram commutes:*

$$\begin{tikzcd} R_\infty \arrow[d, equal] \arrow[r, phantom, "\curvearrowright"] & S_{\infty,J'} \arrow[d, two heads] \\ R_\infty \arrow[r, phantom, "\curvearrowright"] & S_{\infty,J} \end{tikzcd}$$

(5) *The action of R_∞ is compatible with change of level, i.e. the following diagram commutes:*

$$\begin{tikzcd} R_\infty \arrow[d, two heads] \arrow[r, phantom, "\curvearrowright"] & S_{\infty,J} \arrow[d, two heads] \\ R_\emptyset^{\mathrm{univ}} \arrow[r, phantom, "\curvearrowright"] & S_\emptyset/J \end{tikzcd}$$

Proof. Much of this is straightforward, so we'll focus on the trickier aspects, which are the proof of (1) and making the map $\Lambda \to J$ in (4).

(1) Let $d = \operatorname{rank}_{\mathscr{O}} S_\emptyset$. We pick an isomorphism $\mathscr{O}[\Delta_{Q_m}]^d \xrightarrow{\sim} S_{Q_m}$. Modding out by J, we get

$$(\Lambda/J)^d \twoheadrightarrow S_{Q_m}/JS_{Q_m}.$$

and

$$\left(\prod_{\mathbf{Z}_{>0}} \Lambda/J\right)^{\oplus d}_{\mathfrak{p}(\mathscr{F})} \to \left(\prod S_{Q_m}/JS_{Q_m}\right)_{\mathfrak{p}(\mathscr{F})}.$$

We claim that this is an isomorphism; we will check that it is injective and surjective.

For injectivity, suppose you had a tuple of the form

$$((x_m^1),\dots,(x_m^d))\cdot(y_m)^{-1} \mapsto 0$$

where $Z(y_m) \notin \mathscr{F}$. Then there exists $(z_m) \notin \mathfrak{p}$, i.e. $Z(z_m) \notin \mathscr{F}$, such that $(z_m x_m^1,\dots,z_m x_m^d) \mapsto 0$ for all m. This implies $z_m x_m^i = 0$ for all $i, m \gg 0$ (since we know that eventually the levelwise map is injective). We then define

$$z'_m = \begin{cases} z_m & m \gg 0 \\ 0 & \text{otherwise,} \end{cases}$$

so that $(z'_m)(x_m^i) = 0$ for all i. But $Z(z'_m) \notin \mathscr{F}$ because it differs from $Z(z_m)$ by a finite set.

For (4): for $k \gg_J 0$, we have

$$J \cdot \mathbf{T}_\emptyset \supset \mathfrak{m}_\emptyset^k \mathbf{T}_\emptyset$$

and

$$J \cdot S_\emptyset \supset \mathfrak{m}_\emptyset^k S_\emptyset.$$

Let $\mathfrak{m}_\infty \triangleleft R_\infty$ the maximal ideal $(\lambda, X_1, \dots, X_s)$. For any fixed m, $\mathfrak{m}_\infty^k \to 0 \subset \operatorname{End}(S_{Q_m}/J)$ for sufficiently large k. But the tricky issue is that we need to make k uniform in m in order to patch. The point is that the length of S_{Q_m} over Λ/J *is* uniformly bounded, so that lets us choose a k that works uniformly in m.

Now, as above we consider the products

$$\left(\prod_m \Lambda/J\right)_{\mathfrak{p}(\mathscr{F})} = \Lambda/J.$$

and

$$S_{\infty,J} = \left(\prod_m S_{Q_m}/J\right)_{\mathfrak{p}(\mathscr{F})}.$$

For the same ultrafilter $\mathscr{F}$, we get a prime $\mathfrak{p}(\mathscr{F})' \subset R$, and an action

$$\left(\prod_m R_\infty/\mathfrak{m}_\infty^k\right)_{\mathfrak{p}(\mathscr{F})'} \curvearrowright S_{\infty,J}.$$

And the same argument implies

$$R_\infty \twoheadrightarrow R_\infty/\mathfrak{m}_\infty^k \cong \left(\prod_m R_\infty/\mathfrak{m}_\infty^k\right)_{\mathfrak{p}(\mathscr{F})'} \curvearrowright S_{\infty,J}.$$

Since factor-wise we can choose $\Lambda \to R_\infty/\mathfrak{m}_\infty^k$ compatibly with everything, we can then lift it to $\Lambda \to R_\infty$.

This is all compatible with

$$\begin{tikzcd}
R_\emptyset^{\mathrm{univ}}/\mathfrak{m}^k \arrow[r, two heads] \arrow[d, equal] & \mathbf{T}_\emptyset/\mathfrak{m}^k \arrow[d, equal] \arrow[r, phantom, "\curvearrowright"] & S_{\emptyset,J} \arrow[d, equal] \\
(\prod_m R_\emptyset/\mathfrak{m}^k)_{\mathfrak{p}(\mathscr{F})'} \arrow[r, two heads] & (\prod_m \mathbf{T}_\emptyset/\mathfrak{m}^k)_{\mathfrak{p}(\mathscr{F})'} \arrow[r, phantom, "\curvearrowright"] & (\prod_m S_\emptyset/J)_{\mathfrak{p}(\mathscr{F})'}
\end{tikzcd}$$

Putting this together, we get the desired statements. □

8.5. **Completion of the proof.** We will finally prove that $R_\emptyset \xrightarrow{\sim} \mathbf{T}_\emptyset$.

Now define $S_\infty = \varprojlim_J S_{\infty,J}$ – this is finite free over Λ. We have by property (2) in Proposition 8.4.1

$$S_\infty/\mathfrak{a} \xrightarrow{\sim} S_\emptyset.$$

Also, R_∞ acts on S_∞, in a manner commuting with the Λ-action.

$$\begin{tikzcd}
\Lambda \arrow[r] \arrow[d] & R_\infty \arrow[d, two heads] \arrow[rr, phantom, "\curvearrowright"] & & S_\infty \\
\Lambda/\mathfrak{a} \arrow[r] & R_\emptyset^{\mathrm{univ}} \arrow[r, two heads] & \mathbf{T}_\emptyset \arrow[r, phantom, "\curvearrowright"] & S_\emptyset
\end{tikzcd}$$

Now consider $\operatorname{depth}_{R_\infty}(S_\infty)$ – the longest regular sequence. Since the Λ-action on S_∞ factors through R_∞, we obviously have

$$\operatorname{depth}_{R_\infty}(S_\infty) \geq \operatorname{depth}_\Lambda(S_\infty).$$

But S_∞ is finite free over Λ by Proposition 8.4.1 (1), so

$$\operatorname{depth}_\Lambda(S_\infty) = \operatorname{depth}_\Lambda(\Lambda) = s+1$$

Since R_∞ is noetherian local, and S_∞ is finitely generated over R_∞, we can apply the Auslander-Buchbaum theorem, which says that

$$\mathrm{pd}_{R_\infty}(S_\infty)+\mathrm{depth}_{R_\infty}(S_\infty)=\mathrm{depth}_{R_\infty}(R_\infty)=s+1.$$

As we already found that $\mathrm{depth}_{R_\infty}(S_\infty)\geq s+1$, we must have

$$\mathrm{depth}_{R_\infty}(S_\infty)=\mathrm{depth}_{R_\infty}(R_\infty)=s+1$$

and

$$\mathrm{pd}_{R_\infty}(S_\infty)=0,$$

i.e. S_∞ is already projective over R_∞, hence free because R_∞ is a local ring.

Now mod out by $\mathfrak{a}$: we get that $S_\infty/\mathfrak{a}=S_\emptyset$ is free over $R_\infty/\mathfrak{a}R_\infty$. But this action factors throug

$$R_\infty/\mathfrak{a}R_\infty \twoheadrightarrow R_\emptyset \twoheadrightarrow \mathbf{T}_\emptyset \curvearrowright S_\emptyset.$$

So all these maps must also be injective, hence isomorphisms. We finally win.

We used here that R_∞ is regular. We want to show that you can use less commutative algebra to still deduce $R_\emptyset=\mathbf{T}_\emptyset$ (without the freeness of $S_\emptyset$ over $\mathbf{T}_\emptyset$). This is because one doesn't always have good control over (the analogue of) R_∞.

We have

$$s+1=\mathrm{depth}_\Lambda(S_\infty)\leq \mathrm{depth}_{R_\infty}(S_\infty)\leq \dim R_\infty/\mathrm{Ann}_{R_\infty}(S_\infty)\leq \dim R_\infty=s+1.$$

Hence we get equalities everywhere, including

$$\dim R_\infty/\mathrm{Ann}_{R_\infty}(S_\infty)\leq \dim R_\infty.$$

As R_∞ is a domain (obvious here, and will be true more generally), we have $\mathrm{Ann}_{R_\infty}(S_\infty)=0$. Hence $\mathrm{supp}_{R_\infty}(S_\emptyset)=\mathrm{Spec}(R_\infty/\mathfrak{a})$. Hence $\mathrm{Ann}_{R_\infty/\mathfrak{a}}(S_\emptyset)$ is nilpotent. As $R_\infty/\mathfrak{a}\cong R_\emptyset$, we deduce that $\ker(R_\emptyset\to\mathbf{T}_\emptyset)$ is nilpotent. This is enough for applications to modularity, as any closed point factors through the reduced subscheme.

9. BEYOND THE MINIMAL CASE

9.1. **Back to the basic setup.** We had been assuming that $T=\emptyset$. Let's now try to drop this assumption and see where our argument fails. We had arranged that

- $\#T<\infty$, and for all $v\in T$ we have $\#k(v)\equiv \pmod{\ell}$, $\overline{r}|_{G_{F_v}}=1$.
- $\overline{r}|_{G_{F_v}}=1$.
- For $\sigma\in I_{F_v}$, the characteristic polynomial of $\overline{r}(\sigma)$ is $(X-1)^2$, i.e. the action of I_{F_v} is unipotent.

For $v\in T$, we take U_v to be Iw_v instead of $\mathrm{GL}_2(\mathcal{O}_{F,v})$. We use a new version of $R_\emptyset^{\mathrm{univ}}$ that allows deformations of $\overline{r}$ which may ramify at $v\in T$, but forces $\sigma\in I_{F_v}$ to have characteristic polynomial $(X-1)^2$ for $v\in T$.

We introduce analogous version R_Q^{univ}. We find again there are maps

$$\begin{array}{ccc} R_\emptyset^{\mathrm{univ}} & \twoheadrightarrow & \mathbf{T}_\emptyset \\ \uparrow\!\!\uparrow & & \uparrow\!\!\uparrow \\ R_Q^{\mathrm{univ}} & \twoheadrightarrow & \mathbf{T}_Q \end{array}$$

What if we try to run the same argument? Looking at the tangent space, one gets

$$\dim(\mathfrak{m}_Q/(\mathfrak{m}_Q^2,\lambda)) = \dim H^1_{\mathscr{L}^\perp}(G_{F,Q\cup\{T\}\cup\{\ell\}};\mathrm{ad}^0\,\overline{r}(1)) + \#Q + \underbrace{\sum_{v\in T}(\dim L_v - \dim H^0(G_{F_v};\mathrm{ad}^0\,\overline{r}))}_{\text{used to be }0}.$$

Notice the new local term on the right; let's try to calculate its contribution.

We need to describe the condition L_v on a first-order deformation $\sigma \mapsto (1+\phi(\sigma)\epsilon)\overline{r}(\sigma)$, so $\phi \in Z^1(G_{F_v};\mathrm{ad}^0\,\overline{r})$. The condition on I_{F_v} is that the char poly is $(X-1)^2$. This comes out to

$$\det(X\,\mathrm{Id}_2 - \overline{r}(\sigma) - \epsilon\phi(\sigma)\overline{r}(\sigma)) = (X-1)^2.$$

But this turns out to be *automatic* from the condition $\mathrm{Tr}\,\phi(\sigma)=0$.

The upshot is that L_v is all of $H^1(G_{F_v};\mathrm{ad}^0\,\overline{r})$. So the local contribution for each $v \in T$ is

$$(\dim L_v - \dim H^0(G_{F_v};\mathrm{ad}^0\,\overline{r})) = \dim H^0(G_{F_v};\mathrm{ad}^0\,\overline{r}(1)) = 3.$$

Now the argument breaks down, as we're going to get $\Lambda = \mathscr{O}[[S_1,\dots,S_s]]$ and $R_\infty = \mathscr{O}[[X_1,\dots,X_s,Y_1,\dots,Y_{3\#T}]]$. The problem is that R_∞ is too big.

Kisin realized how to deal with this. R_∞ should have dimension $s+1$, but it should be singular. We're getting something too big because we're estimating the tangent space at singular points.

Kisin's insight was that the Galois deformation ring is singular of dimension $s+1$, *but the singularities have a local origin.* He proposes to solve this issue by working *relative* to a local Galois deformation ring.

9.2. **Framed deformation rings.** We want to work with "$R_{\overline{r}|_{G_{F_v}}}^{\mathrm{univ}}$" for $v \in T$. Since the residual representation is trivial, this doesn't exist as a formal scheme (it has too many automorphisms). A way to fix this is to look at framed deformations. We define a "universal lifting ring" $R^\square_{\overline{r}|_{G_{F_v}}}$. Here we do not mod out by conjugation. (We use the word "lifting" to signify that it parametrizes actual homomorphisms, rather than merely "deformations" which parametrize homomorphisms up to conjugation.) So this is bigger by 3 dimensions. In fact, morally one expects $R^\square_{\overline{r}|_{G_{F_v}}}$ to be a power series ring in 3 variables over $R_{\overline{r}|_{G_{F_v}}}^{\mathrm{univ}}$.

9.2.1. *The local lifting ring.* We define a quotient $R_{v,1}$ to parametrize liftings such that for all $\sigma \in I_{F_v}$, the characteristic polynomial is $(x-1)^2$. This is very concrete. Any such representation is automatically tamely ramified, and tame inertia is easy to describe. Such a lifting just amounts to giving $\Phi,\Sigma \in \mathrm{GL}_2(A)$ such that

- $\Phi \equiv \mathrm{Id}_2 \pmod{\mathfrak{m}}$,
- $\Sigma \equiv \mathrm{Id}_2 \pmod{\mathfrak{m}}$,
- $\Phi\Sigma\Phi^{-1} = \Sigma^{\#k(v)}$,
- $\mathrm{char}_\Sigma(X) = (X-1)^2$.
- $\det\Phi$ is fixed.

Let $R_{\mathrm{loc}} = \widehat{\otimes}_{v \in T} R_{v,1}$. We have to pay attention to the difference between liftings and deformations. The global deformation ring doesn't admit a homomorphism from R^{loc} since it has only deformations at T, rather than liftings.

9.2.2. *Global deformation ring framed at T.* To amend this, we introduce a version of the global ring that is framed at T. We define $R_{\emptyset}^{\square_T}$ to represent: pairs $(\rho, \{\alpha_v\})$ such that ρ lifts $\overline{r}$, satisfying previous conditions, and $\{\alpha_v\} \in \ker(\mathrm{GL}_2(A) \to \mathrm{GL}_2(\mathbf{F}))$, modulo an equivalence relation: for any $\beta \in \ker(\mathrm{GL}_2(A) \to \mathrm{GL}_2(\mathbf{F}))$,

$$(\rho, \{\alpha_v\}) \sim (\beta \rho \beta^{-1}, \{\beta \alpha_v\}). \tag{9.2.1}$$

Notice that $\alpha_v^{-1} \rho|_{G_{F_v}} \alpha_v$ is a well-defined homomorphism. So we have a lifting at T, hence we get $R^{\mathrm{loc}} \to R_{\emptyset,1}^{\square_T}$ and $R^{\mathrm{loc}} \to R_{Q,1}^{\square_T}$. (Here the subscript 1 refers to the condition that any $\sigma \in I_{F_v}$ for $v \in T$ has characteristic polynomial $(X-1)^2$, i.e. is unipotent.)

9.2.3. *Framed vs unframed.* What is the relation between $R_{\emptyset}^{\square_T}$ and $R_{\emptyset}$? Suppose you give $r^{\mathrm{univ}} \colon G_F \to \mathrm{GL}_2(R_{\emptyset}^{\mathrm{univ}})$, and you want a framing at T. The framing looks like 4 variables for each v, but you lose one in the end because of the quotient effected by taking β to be a scalar in (9.2.1). So you have a non-canonical presentation

$$R_{Q,1}^{\square_T} \approx R_{Q,1}^{\mathrm{univ}}[[A_1, \ldots, A_{4\#T-1}]]. \tag{9.2.2}$$

9.2.4. *Framed modular forms.* We need to increase our spaces of modular forms to match (9.2.2). So we define

$$S_Q^{\square_T} := S_Q \otimes_{\mathcal{O}} \mathcal{O}[[A_1, \ldots, A_{4\#T-1}]]$$

and

$$\mathbf{T}_Q^{\square_T} := \mathbf{T}_Q[[A_1, \ldots, A_{4\#T-1}]].$$

9.3. **The patching argument.**

We now go through the patching argument in this non-minimal case.

9.3.1. *Adding finite level.* As before, we have a diagram

$$\begin{array}{ccccc} R_1^{\mathrm{loc}} & \longrightarrow & R_{Q,1}^{\square_T} & \curvearrowright & S_{Q,1}^{\square_T} \\ & & & & \\ & & \mathbf{T}_{\emptyset,1}^T & \curvearrowright & S_{\emptyset,1}^T \end{array}$$

for any finite level Q.

The augmentation ideal in this case is

$$\mathfrak{a} = (A_1, \ldots, A_{4\#T-1}, \{\delta - 1\}_{\delta \in \Delta_Q}) \subset \mathcal{O}[[A_1, \ldots, A_{4\#T-1}]][\Delta'_Q]$$

and we have $S_{\emptyset,1}^T = S_{Q,1}^{\square_T}/\mathfrak{a}$.

9.3.2. *Tangent space.* We are interested in the *relative* tangent space of $R^{\square_T}_{Q,1}/R^{\mathrm{loc}}_1$.

Lemma 9.3.1. *Assume $T \neq \emptyset$. Then we have*

$$\dim \mathfrak{m}_{R^{\square_T}_{Q,1}}/((\mathfrak{m}_{R^{\square_T}_{Q,1}})^2, \mathfrak{m}_{R^{\mathrm{loc}}_1}) = \dim H^1_{(\mathscr{L}^T_Q)^\perp}(G_{F,T\cup\{\ell\}\cup Q}, (\mathrm{ad}^0\,\overline{r})(1)) + \#Q + \#T - 1.$$

where $\mathscr{L}^T_Q$ agrees with $\mathscr{L}_Q$ away from T, but is 0 at T (since we are considering things relative R^{loc}_1, what happens locally at T is fixed). Therefore, $(\mathscr{L}^T_Q)^\perp$ is the same as $\mathscr{L}^\perp_Q$ away from T, and at $v \in T$ it's all of $H^1(G_{F_v}, \mathrm{ad}^0\,\overline{r}(1))$.

Proof. Look at Theorem 7.3.3. The new local contribution $\#T - 1$ arises as follows. For $v \in T$, it should be

$$\dim L_v - \underbrace{\dim H^0(G_{F_v}; \mathrm{ad}^0\,\overline{r})}_{=3} + \underbrace{4}_{\text{framing}} = 1.$$

But at the end we have to subtract 1 (if T is non-empty) because of the quotienting by scalar β in (9.2.1). $\square$

9.3.3. *Patching.* We choose Q as in Proposition 7.4.1 to allow a surjection

$$R_\infty \twoheadrightarrow R^{\square_T}_{Q,1}$$

with $R_\infty = R^{\mathrm{loc}}_1[[X_1, \ldots, X_{\#Q+\#T-1}]]$.

We define $\Lambda = \mathscr{O}[[S_1, \ldots, S_{\#Q}, A_1, \ldots, A_{4\#T-1}]]$, which has dimension $4\#T + \#Q$. We let $\mathfrak{a}_\infty = (A_1, \ldots, A_{4\#T-1}, S_1, \ldots, S_{\#Q}) \subset \Lambda$.

Note that

$$\dim R_\infty = \dim R^{\mathrm{loc}}_1 + \#Q + \#T - 1 = (3\#T + 1) + \#Q + \#T - 1 = 4\#T + \#Q.$$

Hence we find that $\dim \Lambda = \dim R_\infty$: the dimensions match again.

$$\begin{tikzcd}
\Lambda \arrow[r, dashed] \arrow[dr, dashed] & R_{\infty,1} \arrow[d, dashed] & & & \\
& R^{\square_T}_{Q,1} \arrow[r, two heads] \arrow[d, two heads] & \mathbf{T}^{\square_T}_{Q,1} \arrow[d, two heads] & \curvearrowright & S^{\square_T}_{Q,1} \arrow[d, two heads] \\
& R^T_{\emptyset,1} \arrow[r, two heads] & \mathbf{T}^T_{\emptyset,1} & \curvearrowright & S^T_{\emptyset,1}
\end{tikzcd}$$

The augmentation $\mathfrak{a} \subset \Lambda$ maps to 0 in $R^{\square_T}_{\emptyset,1}$ and induces

$$S^{\square_T}_{Q,1}/\mathfrak{a} \xrightarrow{\sim} S^T_{\emptyset,1}$$

and $S^T_{Q,1}$ is finite free over $\Lambda/\ker(\Lambda \to \mathscr{O}[[A_1, \ldots, A_{4\#T}]][\Delta_Q])$.

As before, we can patch this to a diagram

$$\begin{tikzcd}
\Lambda \arrow[r, dashed] & R_{\infty,1} \arrow[d, two heads] & & \curvearrowright & S^{\square_T}_{\infty,1} \\
& R^T_{\emptyset,1} \arrow[r, two heads] & \mathbf{T}^T_{\emptyset,1} & \curvearrowright & S^T_{\emptyset,1}
\end{tikzcd}$$

Furthermore,

- $S_{\infty,1}$ is finite free over Λ,
- $\mathfrak{a} \subset \Lambda$ maps to (0) in $R^T_{\emptyset,1}$ and induces

$$S_{\infty,1}/\mathfrak{a} \xrightarrow{\sim} S^T_{\emptyset,1}.$$

Note that $\dim \Lambda = 4\#T + \#Q$. So

$$4\#T + \#Q = \dim R_{\infty,1} \geq \dim R_{\infty,1}/\mathrm{Ann}(S_{\infty,1}) \geq \mathrm{depth}_{R_{\infty,1}}(S_{\infty,1}) \geq \mathrm{depth}_{\Lambda}(S_{\infty,1}) = 4\#T + \#Q.$$

This forces equality to hold everywhere, so $\mathrm{supp}_{R_{\infty,1}}(S_{\infty,1}) = \mathrm{Spec}(R_{\infty,1}/\mathrm{Ann}(S_{\infty,1}))$ is a union of irreducible components of Spec $R_{\infty,1}$. Since this is formally smooth over R_1^{loc} (recall that $R_{\infty,1} = R_1^{\mathrm{loc}}[[x_1, \ldots, X_{\#Q+\#T-1}]]$) it is the pre-image of a union of irreducible components in R_1^{loc}.

By the diagram

$$\begin{tikzcd}
\mathrm{Spec}\, R^{\square_T}_{\emptyset,1} \arrow[d] & \mathrm{supp}_{R^{\square_T}_{\emptyset,1}}(S^T_{\emptyset,1}) \arrow[l, hook'] \arrow[d, equal] & \\
\mathrm{Spec}\, R_{\infty,1}/\mathfrak{a} & \mathrm{supp}_{R_{\infty,1}/\mathfrak{a}}(S^T_{\emptyset,1}) \arrow[l, hook'] \arrow[r, equal] & \mathrm{supp}_{R_{\infty,1}}(S^T_{\emptyset,1}) \cap \mathrm{Spec}\, R_{\infty,1}/\mathfrak{a}
\end{tikzcd}$$

we deduce that $\mathrm{supp}_{R^T_{\emptyset,1}}(S^T_{\emptyset,1})$ is the pre-image of a union of irreducible components of Spec $R_{\mathrm{loc},1}$. But Spec R_1^{loc} has $2^{\#T}$ irreducible components. For each $v \in T$, there are two components: the unramified representations, and unramified twists of the Steinberg ($N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$). So as long as we don't change the local behavior, we're on the same component and we win. Otherwise we're in trouble, and in the next section we'll explain a trick to resolve this issue.

9.3.4. There's a relationship between irreducible components of Spec R_1^{loc} and Spec $R_1^{\mathrm{loc}}/\lambda$, where $\mathfrak{p} \longleftrightarrow \mathfrak{P}$ if $\mathfrak{P} \supset \mathfrak{p}$. In principle this could be many-to-many: generic components could collide over the special fiber, or components could live over the special fiber. In this case, however, this correspondence is 1:1. So it'll suffice to show that $\mathrm{supp}_{R^T_{\emptyset,1}}(S^T_{\emptyset,1})$ contains all components of the special fiber.

9.4. **Deformation rings with nebentypus.** We consider characters of the form

$$\psi = \prod \psi_v \colon \prod_{v \in T} I_v \to \prod_{v \in T} I(F_v^{\mathrm{ab}}/F_v) \xrightarrow{\prod_v \mathrm{Art}_{F_v}^{-1}} \prod_{v \in T} \mathcal{O}_{F,v}^{\times} \twoheadrightarrow \prod_{v \in T} k(v)^{\times} \to \mu_{\ell^\infty}.$$

For any such character, we consider a local lifting ring $R^{\square}_{v,\psi}$ that parametrizes lifts ρ of $\overline{r}|_{G_{F_v}}$ such that if $\sigma \in I_{F_v}$, then

$$\mathrm{char}_{\rho(\sigma)} = (X - \psi_v(\sigma))(x - \psi_v(\sigma)^{-1}).$$

Again this can be written down concretely:

- $\Phi \equiv \mathrm{Id}_2 \pmod{\mathfrak{m}}$,
- $\Sigma \equiv \mathrm{Id}_2 \pmod{\mathfrak{m}}$,

- $\Phi\Sigma\Phi^{-1} = \Sigma^{\# k(v)}$,
- $\mathrm{char}_\Sigma(X) = (X - \psi_v(\sigma))(x - \psi_v(\sigma)^{-1})$.
- $\det \Phi$ is fixed.

As before this is 4-dimensional (relative dimension 3 over $\mathcal{O}$). But the difference is that if $\psi_v \neq 1$, it is *irreducible.*

9.4.1. Since ψ takes values in μ_{ℓ^∞}, it is *trivial mod* ℓ. So we have

$$R^\square_{v,1}/\lambda \cong R^\square_{v,\psi}/\lambda.$$

The picture is: for $R_{v,1}$, we had two generic components reducing to two special components. For $R_{v,\psi}$, we have *one* generic component reducing to (the same!) two special components.

We will develop the story for ψ parallel to the case $\psi = 1$, which we had been considering previously.

9.4.2. *Local deformation rings.* Define as before $R^{\mathrm{loc}}_\psi := \widehat{\bigotimes}_{v\in T} R^\square_{v,\psi}$. We have $\dim R^\square_{v,\psi} = 4$, so $\dim R^{\mathrm{loc}}_\psi = 3\#T + 1$.

9.4.3. *Global deformation rings.* We also have global rings $R^{\mathrm{univ},T}_{Q,\psi}$ and $R^{\square_T}_{Q,\psi}$. A choice of r^{univ} over $R^{\mathrm{univ},T}_{Q,\psi}$ induces an isomorphism

$$R^{\square_T}_{Q,\psi} \xrightarrow{\sim} R^{\mathrm{univ},T}_{Q,\psi}[[A_1,\dots,A_{4\#T-1}]]$$

equipped with $R^{\mathrm{loc}}_\psi \to R^{\square_T}_{Q,\psi}$.

9.4.4. *Modular forms.* Define also $S^T_{Q,\psi} = S_{k,\chi}(U^T_Q,\psi;\mathcal{O})_{\mathfrak{m}}$. This means for $v \in T$, $(U^T_Q) = \mathrm{Iw}_v$ and we look at the forms transforming by the character

$$\mathrm{Iw}_v \twoheadrightarrow k(v)^\times \xrightarrow{\psi_v} \mathcal{O}^\times.$$

We also get a framed version

$$S^{\square_T}_{Q,\psi} := S^T_{Q,\psi} \otimes_{\mathcal{O}} \mathcal{O}[[A_1,\dots,A_{4\#T-1}]].$$

This has an action of

$$\mathbf{T}^{\square_T}_{Q,\psi} := \mathbf{T}^T_{Q,\psi}[[A_1,\dots,A_{4\#T-1}]].$$

9.5. **Patching with nebentypus.**

9.5.1. *Finite level setup.* Let $\Lambda_Q := \mathscr{O}[\Delta_Q][[A_1,\dots,A_{4\#T-1}]]$. We have a diagram

$$\begin{tikzcd} \Lambda_Q \arrow[d] \arrow[dr] & & & \\ R^{\square_T}_{Q,\psi} \arrow[r, two heads] \arrow[d, two heads] & \mathbf{T}^{\square_T}_{Q,\psi} \arrow[d, two heads] & \curvearrowright & S^{\square_T}_{Q,\psi} \arrow[d, two heads] \\ R^{\mathrm{univ},T}_{\emptyset,\psi} \arrow[r, two heads] & \mathbf{T}^{T}_{\emptyset,\psi} & \curvearrowright & S^{T}_{\emptyset,\psi} \end{tikzcd}$$

Furthermore,

- $S^{\square_T}_{Q,\psi}$ is finite free over Λ_Q.
- The augmentation $\mathfrak{a}_Q = (A_1,\dots,A_{4\#T-1},\{\delta-1\}_{\delta\in\Delta_Q})$ maps to 0 in $R^{\mathrm{univ},T}_{\emptyset,\psi}$.
- $S^{T}_{\emptyset,\psi} \cong S^{\square_T}_{Q,\psi}/\mathfrak{a}_Q$.

We can choose Q so that $R^{\mathrm{loc}}_{\psi}[[X_1,\dots,X_{\#Q+\#T-1}]]$ surjects onto each $R^{\square_T}_{Q,\psi}$. Furthermore, we can arrange that all these objects for different ψ coincide mod λ.

9.5.2. *Patching.* Set $R^{\mathrm{loc}}_{\infty,\psi} := R^{\mathrm{loc}}_{\psi}[[X_1,\dots,X_{\#Q+\#T-1}]]$.

Let $\Lambda = \mathscr{O}[[S_1,\dots,S_s,A_1,\dots,A_{4\#T-1}]]$, and $\mathfrak{a} = (S_1,\dots,S_s,A_1,\dots,A_{4\#T-1}) \subset \Lambda$.

The patching argument gives:

$$\begin{tikzcd} \Lambda \arrow[d] & & & \\ R^{\square_T}_{\infty,\psi} \arrow[r, equal] \arrow[d, two heads] & R^{\mathrm{loc}}_{\psi}[[X_1,\dots,X_{s+4\#T-1}]] & \curvearrowright & S^{\square_T}_{\infty,\psi} \arrow[d, two heads] \\ R^{\mathrm{univ},T}_{\emptyset,\psi} \arrow[r, two heads] & \mathbf{T}^{T}_{\emptyset,\psi} & \curvearrowright & S^{T}_{\emptyset,\psi} \end{tikzcd}$$

such that

- $S^{\square_T}_{\infty,\psi}$ is finite free over Λ,
- $S^{\square_T}_{\infty,\psi}/\mathfrak{a} \xrightarrow{\sim} S^{T}_{\emptyset,\psi}$
- $\mathfrak{a} \mapsto (0) \in R^{\mathrm{univ},T}_{\phi,\psi}$.
- These are all independent of ψ mod λ.

Hence the patching argument gives

$$s+4\#T = \mathrm{depth}_{\Lambda}(S^{\square_T}_{\infty,\psi}) \leq \mathrm{depth}_{R^{\square_T}_{\infty,\psi}}(S^{\square_T}_{\infty,\psi}) \leq \dim \frac{R^{\square_T}_{\infty,\psi}}{\mathrm{Ann}_{R^{\square_T}_{\infty,\psi}}(S^{\square_T}_{\infty,\psi})} \leq \dim R^{\square_T}_{\infty,\psi} = s+4\#T$$

therefore

$$\dim(R^{\square_T}_{\infty,\psi}/\operatorname{Ann}_{R^{\square_T}_{\infty,\psi}}(S^{\square_T}_{\infty,\psi})) = \dim R^{\square_T}_{\infty,\psi}.$$

Now, the explicit computation of $R^{\square}_{v,\psi}$ for $\psi_v \neq 1$ shows that Spec $R^{\square}_{v,\psi}$ is irreducible. Some algebra then shows that Spec $R^{\emptyset,\mathrm{loc}}_{\psi}$ is also irreducible. So $\operatorname{supp}_{R^{\square_T}_{\infty,\psi}}(S^{\square_T}_{\infty,\psi}) = \operatorname{Spec} R^{\square_T}_{\infty,\psi}$ for nontrivial ψ. Then

$$\operatorname{supp}_{R^{\square_T}_{\infty,\psi}/\lambda}(S^{\square_T}_{\infty,\psi}/\lambda) = \operatorname{Spec} R^{\square_T}_{\infty,\psi}/\lambda.$$

Now, modulo λ everything is independent of ψ, so once we know it for one ψ, we get it for all of them. In particular, the result for non-trivial ψ implies it for $\psi = 1$.

9.5.3. Recall that the issue in the case $\psi = 1$ had to do with multiple irreducible components. The two components parametrize unramified and multiplicative/semistable representations, respectively. There is a bijection between irreducible components of Spec $R^{\square}_{v,1}$ and Spec $R^{\square}_{v,1}/\lambda$, namely $\mathfrak{p} \longleftrightarrow \mathfrak{B}$ if $\mathfrak{B} \supset \mathfrak{p}$.

$$\begin{tikzcd}
\operatorname{supp}_{R^{\square_T}_{\infty,\psi}/\lambda}(S^{\square_T}_{\infty,\psi}/\lambda) \arrow[r, equal] \arrow[d, hook] & \operatorname{Spec} R^{\square_T}_{\infty,\psi}/\lambda \arrow[d, hook] \\
\operatorname{supp}_{R^{\square_T}_{\infty,1}}(S^{\square_T}_{\infty,1}) \arrow[r, hook] & \operatorname{Spec} R^{\square_T}_{\infty,1}
\end{tikzcd}$$

Now, $S^{\square_T}_{\infty,1}$ is torsion-free over $\mathcal{O}$, so $\operatorname{supp}_{R^{\square_T}_{\infty,\psi}/\lambda}(S^{\square_T}_{\infty,\psi}/\lambda) = \operatorname{Spec} R^{\square_T}_{\infty,\psi}/\lambda$ implies $\operatorname{supp}_{R^{\square_T}_{\infty,1}}(S^{\square_T}_{\infty,1}) = \operatorname{Spec} R^{\square_T}_{\infty,1}$.

Then we get

$$\operatorname{supp}_{R^{\square_T}_{\infty,1}/\mathfrak{a}}(S^{\square_T}_{\infty,1}/\mathfrak{a}) = \operatorname{Spec}(R^{\square_T}_{\infty,1}/\mathfrak{a}).$$

Since the $R^{\square_T}_{\infty,1}/\mathfrak{a}$-action factors through $R^{\mathrm{univ},T}_{\emptyset,1} \twoheadrightarrow \mathbf{T}^T_{\emptyset,1}$, we get

$$\begin{tikzcd}
\operatorname{supp}_{R^{\square_T}_{\infty,1}/\mathfrak{a}}(S^T_{\emptyset,1}) \arrow[d, equal] & & \\
\operatorname{Spec}(R^{\square_T}_{\infty,1}/\mathfrak{a}) & \operatorname{Spec} R^{\mathrm{univ},T}_{\emptyset,1} \arrow[l, hook'] & \operatorname{Spec} \mathbf{T}^T_{\emptyset,1} \arrow[l, hook']
\end{tikzcd}$$

We deduce that $|\operatorname{Spec} R^{\mathrm{univ},T}_{\emptyset,1}| = |\operatorname{Spec}(\mathbf{T}^T_{\emptyset,1})|$ i.e. the kernel of $R^{\mathrm{univ},T}_{\emptyset,1} \to \mathbf{T}^T_{\emptyset,1}$ is nilpotent.

REFERENCES

[FL92] Fontaine, J.; Laffaille, G. Construction de représentations p-adiques. Ann. Sci. École Norm. Sup. 15 (1982), no. 4 547–608.

[Se68] Serre, Jean-Pierre. Abelian l-adic representations and elliptic curves. With the collaboration of Willem Kuyk and John Labute. Revised reprint of the 1968 original. Research Notes in Mathematics, 7. A K Peters, Ltd., Wellesley, MA, 1998. 199 pp. ISBN: 1-56881-077-6