Lecture 1: Overview

Brian Conrad
October 2, 2009
Notes by Sam Lichtenstein

1. Why Study Galois Representations?

Here we discuss the link between analytic L-functions and L-functions attached to varieties.
Let F be a number field and X a smooth projective variety over F.
Define the **partial Euler product**

$$\zeta_X^*(s) = \prod_{v \text{ good place of } F} \zeta(X \bmod v, q_v^{-s}), \qquad \operatorname{Re} s \gg_{\dim X} 0.$$

Here $X \bmod v$ denotes the smooth projective variety over $k(v) \simeq \mathbb{F}_{q_v}$ obtained by reduction of a smooth proper model of X over the valuation ring of F_v.

Example. Let X be an elliptic curve over F. Then

$$\zeta_E^*(s) = \prod_{E \text{ has good reduction at } v} \frac{1 - a_v q_v^{-s} + q_v^{1-2s}}{(1-q_v^{-s})(1-q_v^{1-s})} = \zeta^* F(s)\zeta_F^*(s-1)L^*(E,s)^{-1}, \qquad a_v = q_v + 1 - \#E(k(v)).$$

We would like to do the following.

(1) Fill in the "bad" factors to obtain an L-function with a good functional equation, analytic continuation, etc.
(2) Relate ζ_X^* to arithmetic properties of X. (E.g., Birch-Swinnerton-Dyer conjecture, etc.)

A clue for (1) comes from **Artin representations**:

$$G_F = \operatorname{Gal}(\overline{F}/F) \xrightarrow{\quad\rho\quad} \mathrm{GL}(W) = \text{f.d. v.s./}\mathbb{C}$$
$$\operatorname{Gal}(F'/F) = \text{Gal. gp. of finite Gal. extn.}$$

Note that F'/F is unramified at all but finitely many places. We define the **Artin L-function** of ρ to be

$$L(s,\rho) = \prod_v \det([1 - \rho(\mathrm{Frob}_v) q_v^{-s}]|_{W^{I_v}})^{-1}, \qquad W^{I_v} = \text{subspace of } W \text{ fixed by inertia at } v.$$

Grothendieck gave a related description of ζ_X^* using *continuous* p-adic representations

$$G_F \to \mathrm{GL}(\mathrm{H}^i_{\text{ét},c}(X_{\overline{F}}, \mathbb{Q}_p)) =: \mathrm{GL}(W^i).$$

These are unramified almost everywhere, including at all good places away from p. Here the ith cohomology group W^i above vanishes for $i > 2\dim X$. Grothendieck proved that if we remove the contribution of p-adic places to $\zeta_X^*(s)$ then

$$\zeta_X^*(s) = \prod_i L^*(s, W^i)^{(-1)^i},$$

where $L^*(s, W^i)$ is like the Artin L-function without the "bad" factors and the p-adic places:

$$L^*(s, W^i) = \prod_{\text{good } v \nmid p} \det([1 - \mathrm{Frob}_v\, q_v^{-s}]|_{W^i})^{-1}.$$

Note: The expression for $L^*(s, W^i)$ requires some care, since q_v^{-s} is a *complex number* acting on a *p-adic* vector space. What has to be proved is that the characteristic polynomial for the action of Frob_v on W^i has rational coefficients (and is independent of p), so evaluation using q_v^{-s} makes sense (and the Riemann Hypothesis ensures absolute convergence of the product in a suitable right half-plane depending only in $\dim X$).

We conclude from all this that it is a good idea to study L-functions of reasonable p-adic representations. Representation theory can often be used to fill in the bad factors later.

Eternal dangerous bend: The case of $v|p$ is tricky! The complication is that "unramifiedness" is not the right notion corresponding to "good reduction" for p-adic representations of Galois groups of p-adic fields.

2. Modular Galois representations and modularity lifting theorems

Definitions.

Definition. A **p-adic representation** of G_F is a continuous linear representation $\rho : G_F \to \mathrm{GL}(W)$, where W is a finite dimensional vector space over a p-adic field K (i.e. a finite extension $K/\mathbb{Q}_p$) and ρ is unramified at almost all places v of F.

Example. The repsentation $V_p(E) = T_p(E) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$ arising from the Tate module of an elliptic curve E over F is historically the first really interesting example.

Example. Étale cohomology, with compact support: $W^i = \mathrm{H}^i_{\text{ét},c}(X_{\overline{F}}, \mathbb{Q}_p)$ for any separated F-scheme X of finite type. (Note that this is unramified at all but finitely many places, even if X is not smooth. The proof rests on properties of constructible ℓ-adic sheaves.)

Remark. In the definition of a p-adic representation it is *equivalent* to take the coefficient field to be $\overline{\mathbb{Q}}_p$, because of the following fact: any compact subgroup of $\mathrm{GL}_n(\overline{\mathbb{Q}}_p)$ is contained in $\mathrm{GL}_n(K)$ for a finite extension $K/\mathbb{Q}_p$. The proof of this uses the Baire Category Theorem. [Warning! It is *false* if we consider $\mathbb{C}_p$ instead of $\overline{\mathbb{Q}}_p$!] So we *could* do everything over $\overline{\mathbb{Q}}_p$, but we will find it more convenient to take the coefficient field K to be locally compact.

Definition. A **mod p representation** of G_F is a continuous representation $\overline{\rho} : G_F \to \mathrm{GL}(\overline{W})$, where $\overline{W}$ is a finite dimensional vector space over a finite extension $k/\mathbb{F}_p$. Note that $\mathrm{GL}(\overline{W})$ is thus a discrete topological group, so the continuity condition entails that $\overline{\rho}$ factors through a finite Galois group $\mathrm{Gal}(F'/F)$.

Example. The p-torsion of an elliptic curve: $E[p](\overline{F}) \otimes_{\mathbb{F}_p} \mathbb{F}_{p^r}$.

Example. Étale cohomology: $\mathrm{H}^i_{\text{ét},c}(X_{\overline{F}}, \mathbb{Z}/p\mathbb{Z})$.

Remark. It is "equivalent" to take the coefficient field to be $\overline{\mathbb{F}}_p$.

Reduction of Galois representations.

Proposition. Any p-adic representation $\rho : G_F \to \mathrm{GL}_K(W)$ has a G_F-stable $\mathcal{O}_K$-lattice $\Lambda \subset W$; i.e. ρ induces a map $\overline{\rho} : G_F \to \mathrm{GL}_{\mathcal{O}_K} \Lambda \approx \mathrm{GL}_n(\mathcal{O}_K) \twoheadrightarrow \mathrm{GL}_n(k)$ where $k = \mathcal{O}_K/\mathfrak{m}$. □

(Here by a **lattice** we mean a finitely generated $\mathcal{O}_K$-submodule of W such that $K \otimes_{\mathcal{O}_K} \Lambda = W$.) It is not hard to see that the characteristic polynomial of $\overline{\rho}$ is independent of the choice of lattice Λ.

Theorem (Cor. of Brauer-Nesbit Theorem)**.** *Let $\overline{\rho}^{ss} = \bigoplus\{$Jordan-Holder factors of $\overline{\rho}\}$. Then $\overline{\rho}^{ss}$ has the same characteristic polynomial as $\overline{\rho}$, and is determined up to isomorphism by its characteristic polynomial, and is therefore independent of the choice of Λ.*

In light of the theorem, we shall henceforce call $\overline{\rho}^{ss}$ "the" reduction of ρ. Here are a bunch of things to watch out for:

(1) $\overline{\rho}^{ss}$ is often denoted $\overline{\rho}$, even though it is certainly not just the "reduction mod p" of ρ in general.
(2) $\overline{\rho}^{ss}$ may be unramified at some places where ρ is ramifield. For example, if $\rho(I_v) \subset 1 + \mathfrak{m} \cdot \mathrm{Mat}_n \mathcal{O}_K$, then the inertia at v simply "disappears" mod v.
(3) If $\overline{\rho}^{ss}$ is irreducible, then in fact the only stable lattices in W were of the form $\pi^i \Lambda$, where π is a uniformizer for K and $i \in \mathbb{Z}$.
(4) Irreducibility is *not* the same as *absolute* irreducibility = irreducibility over $\overline{k}$.
(5) ρ might be absolutely irreducible over K, yet $\overline{\rho}^{ss}$ could be not only reducible, but even completely trivial! (Hence completely reducible...)

Exercise: If ρ is reducible then any Jordan-Holder filtration of ρ induces a similar filtration for $\overline{\rho}^{ss}$. So the last warning above is "one-directional".

Modular Galois representations. Let $f \in S_k(\Gamma_1(N), \chi)$ be a Hecke eigenform of weight $k \geq 1$. Let $K_f \subset \mathbb{C}$ be the field generated over $\mathbb{Q}$ by all the Fourier coefficients $a_\ell(f)$ of f for primes $\ell \nmid N$. Then K_f is a number field containing the values of the Nebentypus χ. Let λ be a place of K_f lying over p.

Theorem (Deligne, Deligne-Serre, Ribet)**.** *There exists a unique continuous irreducible p-adic representaion*

$$\rho_{f,\lambda} : G_{\mathbb{Q}} \to \mathrm{GL}_2(K_{f,\lambda})$$

unramified at all $\ell \nmid Np$, such that for all such ℓ we have

$$\textit{sum of eigenvalues of } \mathrm{Frob}_\ell = \mathrm{Tr}(\rho_{f,\lambda}\,\mathrm{Frob}_\ell) = a_\ell(f) [= T(\ell)\textit{-eigenvalue of } f]$$

and

$$\det \circ \rho_{f,\lambda} = \chi \cdot \epsilon_p^{k-1}$$

where $\epsilon_p : G_{\mathbb{Q}} \to \mathbb{Z}_p^\times \subset \mathcal{O}_{K_f,\lambda}^\times$ is the p-adic cyclotomic character.

In particular, for $\ell \nmid Np$, the characteristic polynomial of $\rho_{f,\lambda}(\mathrm{Frob}_\ell)$ is

$$t^2 - a_\ell(f)t + \chi(\ell)\ell^{k-1} \in K_f[t] \subset K_{f,\lambda}[t],$$

a non-obvious integrality property. Note that this polynomial is independent of λ.

Remark. The independence of λ and the precise control on the unramified primes implies that the collection $\{\rho_{f,\lambda}\}_\lambda$ is a "compatible" family of of representations, with respect to K_f-characteristic polynomials, just like $\{V_pE\}_p$ is a "compatible" family of represention with respect to $\mathbb{Q}$-characteristic polynomials. Cf. Serre's book *Abelian ℓ-adic representations*.

Let us look at the partial Artin L-functions

$$L^*(s, \rho_{f,\lambda}) = \prod_{\ell \nmid Np} \det\left(1 - \rho_{f,\lambda}(\mathrm{Frob}_\ell) \cdot \ell^{-s}\right)^{-1} = \prod_{\ell \nmid Np} \frac{1}{1 - a_\ell(f)\ell^{-s} + \chi(\ell)\ell^{k-1-2s}} =: L^*(s, f).$$

Remark. Note that for a complex conjugation c, $\det \rho_{f,\lambda}(c) = \chi(-1)\epsilon_p^{k-1} = (-1)^k(-1)^{k-1} = -1$, so all the representations produced by the theorem above are *odd*!

Now consider the (semisimplified) reduction $\overline{\rho}_{f,\lambda} : G_{\mathbb{Q}} \to \mathrm{GL}_2(k_{f,\lambda})$, which is continuous and semisimple, but might be reducible. In general, we say a mod-p representation $\overline{\rho}$ **is modular** if it is isomorphic over $\overline{\mathbb{F}}_p$ to some $\overline{\rho}_{f,\lambda}$.

Just *suppose* $\overline{\rho}_{f,\lambda}$ happens to be absolutely irreducible. By the last remark, it, too, is odd. Serre's conjecture is concerned with when mod-p representations with such properties are in fact modular.

Note that $\overline{\rho}_{f,\lambda}$ does not determine k or N. There could be congruences "$g \equiv f$" modulo λ for some eigenform $g \in S_{k'}(\Gamma_1(N'), \chi')$ (with the congruence taken in the sense of Fourier coefficients, say, relative to a p-adic place of $\overline{\mathbb{Q}}$ over λ on K_f and some chosen p-adic place of K_g). This would imply that $\overline{\rho}_{f,\lambda} = \overline{\rho}_{g,\lambda}$. This is actually abusive notation, since to obtain such a comparison, we might need to extend scalars on the residue fields of these reductions.

Wiles's insight. The prototype of a **modularity lifting theorem** is the following.

Theorem (Not really a theorem)**.** *Given any p-adic representation $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$ such that $\overline{\rho}$ is irreducible and modular over $\overline{\mathbb{F}}_p$, and ρ is "nice" (at p, in the sense of p-adic Hodge theory!) then ρ is modular.*

In this seminar, we'll focus on those ρ such that

$$\rho|_{D_p} \approx \begin{pmatrix} \psi_1 & * \\ 0 & \psi_2 \end{pmatrix}$$

where D_p is the decomposition group at p, ψ_2 is an unramified character, and ψ_1 is ϵ_p^{k-1} times an unramified character. These representations are "essentially like the ones that come from elliptic curves with good ordinary reduction at p".

3. Applications of the method

- Serre's conjecture.
- Sato-Tate.
- Gross-Zagier, Heegner points, Kolyvagin (need to provide a finite map $X_0(N_E) \to E$ over $\mathbb{Q}$, which is done via Faltings' theorem and the "modularity" of $V_\ell(E)$).
- FLT. (Modularity of the Galois rep. attached to the Frey curve.)

Lecture 2: Serre's conjecture and more

Akshay
October 9, 2009
Notes by Sam Lichtenstein

Fix embeddings $\overline{\mathbb{Q}} \hookrightarrow \mathbb{C}$ and $\overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$, and let k denote a finite subfield of the residue field of $\overline{\mathbb{Q}}_p$.

1. Serre's conjecture

Here's the conjecture:

Let $\overline{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ be irreducible and odd. Then there exists a newform f whose Galois representation $\rho_f : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$ satisfies $\overline{\rho}_f \cong \overline{\rho}$. (Here $\overline{\rho}_f$ always means semisimplication!) Moreover f is of level $N(\overline{\rho})$ and weight $k(\overline{\rho})$ to be discussed below.

Remark. Apropos of reduction mod p: If V is a $\mathbb{Q}_p$-vector space and $G \subset \mathrm{GL}(V)$ is a compact subgroup, then there exists a G-fixed lattice in V for the following reason. Pick any lattice $L \subset V$. Then the G-stabilizer of L is open and of finite index. So $\Lambda = \sum_{g \in G} gL \subset V$ is also a lattice, and it is definitely G-stable. The same works with coefficients in any finite extension of $\mathbb{Q}_p$, or even in $\overline{\mathbb{Q}}_p$ (since we saw last time that in this latter case the image is contained in $\mathrm{GL}_n(K)$ for some subfield K of finite degree over $\mathbb{Q}_p$.

The level $N(\overline{\rho})$. Serre conjectured that $N(\overline{\rho})$ = Artin conductor of $\overline{\rho}$, which has the following properties.

- $(p, N(\overline{\rho})) = 1$.
- For $\ell \neq p$, the ℓ-adic valuation $\mathrm{ord}_\ell N(\overline{\rho})$ depends only on $\overline{\rho}|_{I_\ell}$, and is given by

$$\mathrm{ord}_\ell N(\overline{\rho}) = \sum_{j \geq 0} \frac{1}{[G_0 : G_j]} \dim(V/V^{G_j})$$

  Here, we set $K = \overline{\mathbb{Q}}^{\ker \overline{\rho}}$ to be the the field cut out by $\overline{\rho}$, and G_j to be image under $\overline{\rho}$ of the lower-numbered ramification filtration at ℓ of $\mathrm{Gal}(K/\mathbb{Q})$. In other words, if w is a place of K over ℓ, then

$$G_j = \overline{\rho}\{\sigma \in I_\ell \mid \mathrm{ord}_w(\sigma x - x) > j, \forall x \in \mathcal{O}_{K,w}\}.$$

  The filtration goes

$$G_0 = \overline{\rho}(I_\ell) \supset G_1 \supset G_2 \supset \cdots$$

  The first step is of index prime to ℓ, while the latter groups are all ℓ-groups. If K is tamely ramified or unramified at ℓ, then $\mathrm{ord}_\ell N(\overline{\rho}) = \dim(V/V^{I_\ell})$. The Hasse–Arf theorem ensures that the proposed formula for the ℓ-adic ordinal above is actually an integer.

The weight $k(\overline{\rho})$.

Theorem (Deligne)**.** *Suppose f is a newform of weight $< p$ and level prime to p (so χ_f is unramified at p). Suppose f is **ordinary** at p, meaning $a_p(f) \in \overline{\mathbb{Z}}_p^\times$. Then $\overline{\rho}_f$ has a unique 1-dimensional **unramified** quotient; i.e.*

$$\overline{\rho}_f|_{D_p} \sim \begin{pmatrix} \alpha\omega^{k-1} & * \\ 0 & \beta \end{pmatrix}$$

*for **unramified** characters $\alpha, \beta : D_p \to \overline{\mathbb{F}}_p^\times$ and ω the mod-p cyclotomic character.*

It follows that

$$\overline{\rho}_f|_{I_p} \sim \begin{pmatrix} \omega^{k-1} & * \\ 0 & 1 \end{pmatrix}.$$

This can be seen concretely in the case of elliptic curves E with ordinary reduction: for $\rho_f = V_\ell(E)$ the "connected-étale sequence"

$$\mathcal{E}[p^n]^0 \to \mathcal{E}[p^n] \to \mathcal{E}[p^n]/\mathcal{E}[p^n]^0$$

associated to the p^n-torsion on the Néron model $\mathcal{E}$ has last quotient is unramified. Now take limits on generic fibers to deduce the theorem in this case.

Serre conjectured that

$$k(\overline{\rho}) := \begin{cases} 1+pa+b & \text{“most of the time”} \\ 1+pa+b+p-1 & \text{...} \end{cases}$$

is the minimal weight at prime-to-p level. Here $a \leq b$ are integers to be defined below. In the ordinary, low ($< p$) weight case, $a = 0, b = k-1$.

We need to detour into the structure of $I = I_p \subset G_{\mathbb{Q}}$. By definition $I_w \lhd I \twoheadrightarrow I_t$, where I_w, the wild ramification group, is the largest pro-p subgroup.

Proposition. $I_t \cong \mathrm{Hom}(\mathbb{Q}/\mathbb{Z}, \overline{\mathbb{F}}_p^\times) = \varprojlim_r \mathbb{F}_{p^r}^\times = \prod_{\ell \neq p} \mathbb{Z}_\ell(1)$. □

Think: $\widehat{\mathbb{Z}}$ minus the p-part. The Tate-twisting notation records how the canonical Frobenius element in D_p/I_p acts on the abelian quotient I_t of I_p. The map from left to right is $g \mapsto g(\theta_r)/\theta_r$ where $\theta_r^{p-1} = p$. The action of $\mathrm{Frob}_p \in D_p/I_p$ is by raising to the pth power on the right side. The composite quotient map

$$\psi_r : I_t \twoheadrightarrow \mathbb{F}_{p^r}^\times$$

is called the *level-r fundamental character*, though the more canonical collection is its p-powers (thereby being “independent of the choice of $\overline{\mathbb{F}}_p$”).

We can deduce that

$$(\overline{\rho}|_{I_p})^{\mathrm{ss}} \cong \left(\begin{smallmatrix} \chi_1 & 0 \\ 0 & \chi_2 \end{smallmatrix}\right).$$

To see this, note that $\overline{\rho}$ is assumed irreducible. On one hand I_w is pro-p, so by a counting argument it must fix a nontrivial subspace when acting on a vector space over a finite field of characeristic p. On the other hand I_p/I_w is abelian, so it has no irreducible 2-dimensional representations. Hence $\overline{\rho}_{I_p}$ is not itself irreducible; i.e. it is upper triangular, so its semisimplification splits as a direct sum of characters.

Now since $\overline{\rho}|_{I_p}$ extends to a representation of D_p, the pair $\{\chi_1, \chi_2\}$ must be preserved under the Frobenius action of D_p/I_p. In other words, we have

$$\begin{cases} \chi_1^p = \chi_1 \\ \chi_2^p = \chi_2 \end{cases} \qquad \text{or} \qquad \begin{cases} \chi_1^p = \chi_2 & \chi_1^{p^2} = \chi_1 \\ \chi_2^p = \chi_1 & \chi_2^{p^2} = \chi_1 \end{cases}$$

In the first case, each, χ_i factors through $I_t \to \mathbb{F}_p^\times$. In the second case, each χ_i factors through $I_t \to \mathbb{F}_{p^2}^\times$.

So in the first case we can write $\chi_1 = \omega^a, \chi_2 = \omega^b$ for $0 \leq a \leq b$, where $\omega : I_t \to \varprojlim \mathbb{F}_{p^r}^\times \twoheadrightarrow \mathbb{F}_p^\times$ is the mod-p cyclotomic character. In the second case we can likewise write $\chi_1 = \psi^{a+pb}$, $\chi_2 = \psi^{pa+b}$ where $\psi : I_t \to \mathbb{F}_{p^2}^\times$ is the level-2 fundamental character. These are the a, b in Serre's conjecture.

The exceptional case $k(\overline{\rho}) = 1+pa+b+p-1$. Now we address where this case comes from (but without precisely defining it). Consider the special cases

$$\overline{\rho}|_{I_p} \sim \left(\begin{smallmatrix} \omega^2 & * \\ 0 & 1 \end{smallmatrix}\right)$$

and

$$\overline{\rho}|_{I_p} \sim \left(\begin{smallmatrix} \omega & * \\ 0 & 1 \end{smallmatrix}\right).$$

In the first case the guess is $k(\overline{\rho}) = 3$. In the second case the “standard” guess ($a = 0, b = 1$) is $k(\overline{\rho}) = 2$. But a naive combinatorial estimate says that the number of representations of the second type is roughly twice as much as the number of the first type. On the other hand these are certainly fewer modular forms of weight 2 than of weight 3. The “corrected” guess of $p+1$ for the second case when $a = 0$ and $b = 1$ could provide the necessary extra modular representations.

Note: $\overline{\rho}|_{D_p}$ “comes from” a finite flat group scheme over $\mathbb{Z}_p$ if it arises in weight 2; this property depends only on the restriction to inertia, and it can be characterized in purely Galois-theoretic terms. This leads to a special case in Serre's conjecture related to the case $k(\overline{\rho}) = 2$.

Emerton on Serre's conjecture. Matt Emerton has a version of "mod p local Langlands" which gives the following picture. There is a natural action of $\mathrm{GL}_2(\mathbb{A}_f)$ (with $\mathbb{A}_f$ the finite adeles) on

$$\mathrm{Hom}_{G_{\mathbb{Q}}}(\overline{\rho}, \varinjlim_N \mathrm{H}^1(X(N), \overline{\mathbb{F}}_p)) \cong \bigotimes_q{}' \pi_q(\overline{\rho}),$$

where the right side is a "factorization" into local "mod p automorphic" representations. Here $\pi_q(\overline{\rho})$ is finite length but not necessarily irreducible, and depends only on $\overline{\rho}|_{D_q}$. Supose $\overline{\rho} = \overline{\rho}_f$ for $f \in S_k^{\mathrm{new}}(N)$. Then in fact

$$\overline{\rho} \hookrightarrow \mathrm{H}^1(X(N), \mathrm{Sym}^{k-2} \overline{\mathbb{F}}_p^2).$$

Here $\mathrm{Sym}^{k-2} \overline{\mathbb{F}}_p^2$ is viewed as a local system on $X(N)$ as the Tate module of the "universal elliptic curve" (up to some subtleties at the cusps). The right side is almost the same as [*need to clarify appearance of* $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$*-invariants below*]

$$(\mathrm{H}^1(X(N(\rho)), \overline{\mathbb{F}}_p) \otimes \mathrm{Sym}^{k-2} \overline{\mathbb{F}}_p^2)^{\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})} = ((\bigotimes_q{}' \pi_q(\overline{\rho}))^{k(Np)} \otimes \mathrm{Sym}^{k-2} \overline{\mathbb{F}}_p^2)^{\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})} \neq 0$$

[*This needs to be extended a bit more to explain the relation with "independence" of the* N *and the* k *in Serre's conjecture.*]

2. Hecke algebras

Let $V = S_2(\Gamma_0(N))$ for N squarefree. Let $\mathbb{T} \subset \mathrm{End}(V)$ be the $\mathbb{Z}$-subalgebra generated by all Hecke operators $T(p), p \nmid N$ and $U_p, p|N$. (Recall that $U_p : \sum a_n q^n \mapsto \sum a_{np} q^n$.)

Fact: $\mathbb{T}$ is finite over $\mathbb{Z}$.

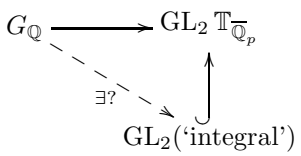
Proof. One approach is to show that $\mathbb{T}$ preserves a lattice in V, by using the arithmetic theory of modular curves (with models over $\mathbb{Z}$). An alternative which is easier to carry out rigorously and involves just topological/analytic tools is to embed V into $\mathrm{H}^1(X_0(N), \mathbb{C})$ and extend the $\mathbb{T}$-action to this space and prove it preserves the lattice of integral cohomology (which can also be studied in terms of group cohomology). This will be addressed in all weights ≥ 2 in Baran's later lecture. □

Fact: The natural map from $\mathbb{T}_{\mathbb{C}} := \mathbb{T} \otimes_{\mathbb{Z}} \mathbb{C}$ onto the subalgebra $\mathbb{C}[T(p), U_p \mid p \in \mathbb{Z}] \subset \mathrm{End}(V)$ is an isomorphism; that is, $\mathbb{T}_{\mathbb{C}}$ acts faithfully on V. This will also be proved in Baran's lecture (in any weight at least 2).

Fact: V is a free $\mathbb{T}_{\mathbb{C}}$ module of rank 1.

Proof. It is enough to construct a cyclic vector f; i.e., $T \mapsto Tf$ gives a surjection $\mathbb{T}_{\mathbb{C}} \twoheadrightarrow V$. (It is automatically then injective since $\mathbb{T}$ acts faithfully on V.) By multiplicity 1, we have $V = \bigoplus_{\text{newforms } f_i} V_i$ where V_i is the generalized Hecke eigenspace corresponding to f_i. It suffices to check the existence of a cyclic vector for each V_i, due to the Chinese Remainder Theorem for coprime maximal ideals of $\mathbb{T}_{\mathbb{C}}$ (which corresponding to eigenforms). The existence of a cyclic vector for each V_i can be done explicitly. □

By the last fact, $\mathrm{H}^1(X_0(N), \mathbb{C}) \cong V \oplus \overline{V}$ is free of rank 2 over $\mathbb{T}_{\mathbb{C}}$. Consequently $\mathrm{H}^1(X_0(N), \overline{\mathbb{Q}}_p)$ is free of rank 2 over $\mathbb{T}_{\overline{\mathbb{Q}}_p}$. The latter is $\mathbb{T}_{\overline{\mathbb{Q}}_p}$-linearly isomorphic to $\mathrm{H}^1_{\text{ét}}(X_0(N)_{\overline{\mathbb{Q}}}, \overline{\mathbb{Q}}_p)$, which also has a $G_{\mathbb{Q}}$-action (that is Hecke equivariant, due to an alternative way to define the Hecke action via correspondences between modular curves over $\mathbb{Q}$). So we obtain a "modular" Galois representation:



We'd like to produce a $G_{\mathbb{Q}}$-stable $\mathbb{T}_{\overline{\mathbb{Z}}_p}$-lattice inside our rank 2 $\mathbb{T}_{\overline{\mathbb{Q}}_p}$ module. This approach gets involved with delicate commutative algebra properties of integral Hecke algebras (Gorenstein condition, etc.), and in more general settings it is simpler to bypass such subtleties at the outset. So we will use a slicker method with wider applicability which avoids making such a Hecke lattice.

Example. Consider level $N = 33$. Then $\dim(S_2) = 3$. The cusp forms in question come from two elliptic curves. The first $y^2 + y = x^3 \pm x^2$ has conductor 11, giving rise to

$$f = q\prod_n (1-q^n)^2(1-q^{11n})^2 = q - 2q^2 - q^3 + 2q^4 + q^5 \pm 2q^6$$

of level 11, hence $f'(z) := f(3z)$ is level 33. The second $y^2 + xy = x^3 + x^2 - 11x$ gives rise to $g = q + q^2 + q^3 - q^4 - 2q^5 \pm 2q^6$ in level 33. Observe that $f \equiv g \bmod 3$, which is no accident. Indeed, the Hecke algebra $\mathbb{T}$ acting on the lattice $\mathbb{Z}f \oplus \mathbb{Z}f' \oplus \mathbb{Z}g$ in S_2 is generated over $\mathbb{Z}$ by U_3, which acts by

$$g \mapsto -g, f' \mapsto f, f \mapsto -f - 3f'.$$

From this we can find

$$\mathbb{T} \cong \mathbb{Z}[x]/(x+1)(x^2+x+3).$$

So $\operatorname{Spec}\mathbb{T}$ lying over $\operatorname{Spec}\mathbb{Z}$ has two irreducible components,

$$\operatorname{Spec}\mathbb{Z} = \operatorname{Spec}\mathbb{Z}[x]/(x+1),\ \ \operatorname{Spec}\mathbb{Z}[x]/(x^2+x+3),$$

which happen to meet at the fiber over $(3) \in \operatorname{Spec}\mathbb{Z}$. (This is precisely the reason for the congruence observed earlier, as we will see in a moment.) The fiber in question consists of a single maximal ideal $\mathfrak{m} \in \operatorname{Spec}\mathbb{T}$, the kernel of

$$\mathbb{T} \overset{\text{act on } \mathbb{Z}f}{\to} \mathbb{Z} \twoheadrightarrow \mathbb{F}_3.$$

If we consider the completed localization $\mathbb{T}_\mathfrak{m}$ then we claim that after a suitable conjugation, $G_\mathbb{Q} \to \mathrm{GL}_2(\mathbb{T}_{\overline{\mathbb{Q}}_3})$ factors through $\mathrm{GL}_2(\mathbb{T}_\mathfrak{m})$. Once this is done, then using the two specializations $\mathbb{T}_\mathfrak{m} \to \mathbb{Z}_3$ corresponding to the two elliptic curves then recovers the 3-adic Tate modules of these elliptic curves as deformations of a common mod-3 residual representation.

But how to make the representation land in $\mathrm{GL}_2(\mathbb{T}_\mathfrak{m})$? Consider the 3-adic eigenforms associated to minimal primes of $\mathbb{T}$ below $\mathfrak{m}$, of which there are 2 and so actually the ones from the elliptic curves above (for a unique prime over 3 in the quadratic field associated to the second component of $\mathbb{T}$). This gives representations from $G_\mathbb{Q}$ into $\mathrm{GL}_2(\mathbb{Z}_3)$ which are conjugate modulo 3. One checks that these mod-3 representations are irreducible, and hence absolutely irreducible (due to oddness). Thus, the *local* fiber product ring

$$R = \mathbb{Z}_3 \times_{\mathbb{F}_3} \mathbb{Z}_3 = \{(a,b) \in \mathbb{Z}_3 \times \mathbb{Z}_3 \,|\, a \equiv b \bmod 3\}$$

contains $S = \mathbb{T}_\mathfrak{m}$ and we get a representation $G_\mathbb{Q} \to \mathrm{GL}_2(R)$ upon fixing an isomorphism of the mod-3 reductions. Note that the traces in R at Frobenius elements away from 3 and 11 all lie in S, since $T_\ell \in \mathbb{T}$ "is" the trace (as can be checked modulo each minimal prime of the reduced $\mathbb{T}_{\mathbb{Q}_3}$). This is the key to descending the representation into $\mathrm{GL}_2(S)$, as we explain next.

3. Descent for Galois representations

Let R be a complete local ring with maximal ideal $\mathfrak{m}_R$. Let $\rho : G_\mathbb{Q} \to \mathrm{GL}_n(R)$ be residually absolutely irreducible and continuous. Suppose further more that ρ is odd. Let S be a complete local subring of R with local inclusion map, so $\mathfrak{m}_S = \mathfrak{m}_R \cap S$ and we get an induced isomorphism of residue fields $S/\mathfrak{m}_S \cong R/\mathfrak{m}_R$. Assume that $\operatorname{tr}\rho(g) \in S$ for all $g \in G_\mathbb{Q}$.

Theorem. *If $n = 2$ and the residue characteristic is not 2 then some $\mathrm{GL}_2(R)$-conjugate of ρ is valued in $\mathrm{GL}_2(S)$.*

Proof. The argument is elementary, and apparently due to Wiles. By oddness, we can assume $\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right) \in \operatorname{im}\rho$. For any $\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right) \in \operatorname{im}\rho$, the trace $2a = \operatorname{tr}(\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right) + \left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right))$ lies in S, so $a \in S$. Similarly one finds $d \in S$. By residual irreducibility there is $g \in G_\mathbb{Q}$ with $\rho(g) \sim \left(\begin{smallmatrix}*&u\\*&\end{smallmatrix}\right)$ where u is an R-unit. Conjugate by $\left(\begin{smallmatrix}u&0\\0&1\end{smallmatrix}\right)$, and we find that $\rho(g) \sim \left(\begin{smallmatrix}*&1\\*&*\end{smallmatrix}\right)$ for some g. Messing around with this and the previous idea, one can conclude that $b, c \in S$ as well. □

Note that the preceding argument did not use the completeness of S. Now we use it. [*Where do we ever use completeness of R or S below?*] Taking S and R as above, and imposing no hypotheses on n or the residue characteristic, we have:

Theorem. *Assume $\rho : G \to \mathrm{GL}_n(R)$ is residually absolutely irreducible, where G is any group at all. Then some $\mathrm{GL}_n(R)$-conjugate of ρ is valued in $\mathrm{GL}_n(S)$.*

Proof. By Jacobson Density and the residual absolute irreducibility of ρ, there exist

$$x_1, \ldots, x_{n^2} \in \rho(G) \subset M_n(R)$$

such that $\overline{x}_i$ span $M_n(k)$, where $k = R/\mathfrak{m}_R$ is the residue field. It follows that the x_i's themselves freely span $M_n(R)$. (Relate them to a basis by a matrix; the reduction of that matrix mod $\mathfrak{m}_R$ is invertible over k, so it must be invertible over R itself.)

Let B be the S-submodule of $M_n(R)$ freely spanned over S by the x_i. It is free of rank n^2. The claim is that B is in fact an S-algebra containing $\rho(G)$. To see this, take $y \in \rho(G)$. We can write $y = \sum a_i x_i$ for $a_i \in R$. For each $1 \leq j \leq n^2$, the trace $\operatorname{tr}(yx_j)$ is equal to $\sum_i a_i \operatorname{tr}(x_i x_j)$. Consider the matrix

$$(\operatorname{tr}(x_i x_j)) \in M_{n^2} S.$$

Due to non-degeneracy of the trace pairing for matrix algebras over a field, a matrix of traces of products of basis elements for a matrix algebra over a field is invertible. So the reduction of this matrix mod $\mathfrak{m}_R$ (the same as its reduction mod $\mathfrak{m}_S$) is invertible. Hence it is invertible itself, so the a_i are in S and hence $y \in B$. Thus, $\rho(G) \subset B$. In particular $1 \in B$. It's not hard to check B is closed under multiplication, so it's a finite S-algebra that is free of rank n^2 and contains $M_n(S)$.

If k' denotes the residue field of S, then since the map $M_n(S) \to M_n(R)$ induces the injective map $M_n(k') \to M_n(k)$ modulo maximal ideals we conclude that the inclusion $M_n(S) \to B$ induces an injective map $M_n(k') \to B \otimes_S k'$. But $B \otimes_S k'$ has rank n^2, so $M_n(S) \to B$ is a map between finite free S-modules of rank n^2 and induces an isomorphism modulo $\mathfrak{m}_S$. Thus, it is an equality. □

4. Universal deformation ring

As before let k be a finite field and $\overline{\rho} : G \to \mathrm{GL}_n(k)$ an absolutely irreducible representation of a profinite group G. A **lifting** of $\overline{\rho}$ over a complete local Noetherian ring A with residue field k is a representations $\rho : G \to \mathrm{GL}_n(A)$ equipped with an isomorphism $\rho \otimes_A k \cong \overline{\rho}$. We will be especially interested in the case when $G = G_{\mathbb{Q},S}$, the Galois group of the largest extension of $\mathbb{Q}$ unramified outside of a fixed finite set of places S, or when G is the Galois group of a local (especially p-adic) field. These groups satisfy a certain finiteness property Φ_p: their open subgroups have only finitely many index-p open subgroups.

Claim. Assume that G satisfies Φ_p. There exists a complete local noetherian ring $R_{\overline{\rho}}$ and a deformation $\rho_{\text{univ}} : G_{\mathbb{Q},S} \to \mathrm{GL}_n(R_{\overline{\rho}})$ such that for any deformation (ρ_A, A) there exists a unique ring map $R_{\overline{\rho}} \to A$ such that ρ_A factors through ρ_{univ}, up to residually trivial conjugation. (Here the map $\mathrm{GL}_n(R_{\overline{\rho}}) \to \mathrm{GL}_n)(A)$ is induced by the map $R_{\overline{\rho}} \to A$.)

The proof of this will be explained next time by Mok.

Example. Let G be a finite group of order not divisible by p and consider $G \xrightarrow{\overline{\rho}} \mathrm{GL}_n(k)$ where the characteristic of k is p. Then $R_{\overline{\rho}} = W(k)$, the ring of Witt vectors for k. This will follow from the vanishing of p-torsion group cohomology for G and the computation of the "reduced" cotangent space to the deformation ring as in Mok's talk next time.

Example. Suppose $\overline{\rho} : G_{\mathbb{Q},S} \to \mathrm{GL}_2(k)$ is odd, and $\mathrm{H}^2(G_{\mathbb{Q}}, \mathrm{Ad}^0(\overline{\rho})) = 0$. Then $R_{\overline{\rho}} = W(k)[\![X_1, X_2, X_3]\!]$. So generically, one expects the universal deformation ring to be 3-dimensional over $W(k)$.

5. Hecke algebras again

Let $\overline{\rho} : G_{\mathbb{Q},S} \to \mathrm{GL}_2(k)$ be absolutely irreducible. Pick a level N. Let $f_1, \ldots, f_m$ be all the newforms of weight 2 and level dividing N, such that $\overline{\rho}_f \sim \overline{\rho} \otimes_k \overline{\mathbb{F}}_p$; we assume this set of f_i's is non-empty! Let f_i have coefficients contained in K_i, a number field with maximal order $\mathcal{O}_i$, and let $\mathcal{O}_{i,\lambda}$ be the completion of $\mathcal{O}_i$ in $\overline{\mathbb{Q}}_p$.

Let $\mathbb{T}$ be the $W(k)$-subalgebra of $\prod \mathcal{O}_i$ spanned by the images of all the $T(\ell)$ with $(\ell, Np) = 1$.

We have a map $\mathbb{T} \to \mathcal{O}_{i,\lambda} \to \overline{\mathbb{F}}_p$ sending $T(\ell)$ to $\operatorname{tr} \rho(\mathrm{Frob}_\ell)$, independent of i. Call the kernel $\mathfrak{m} \subset \mathbb{T}$, and let $\mathbb{T}_{\mathfrak{m}}$ be the completed localization. Thus, the representation

$$\prod \rho_{f_i} : G_{\mathbb{Q},S} \to \mathrm{GL}_2(\prod \mathcal{O}_{i,\lambda})$$

admits a conjugate valued in $\mathrm{GL}_2(\mathbb{T}_m)$, by using the same kind of argument carried out earlier with the elliptic curves of levels 11 and 33. Note that the residue field of $\mathbb{T}_{\mathfrak{m}}$ is equal to k.

By universality of $R_{\overline{\rho}}$ we obtain a local $W(k)$-algebra map $R_{\overline{\rho}} \twoheadrightarrow \mathbb{T}_{\mathfrak{m}}$ satisfying $\operatorname{tr}\rho\,\mathrm{Frob}_\ell \mapsto T(\ell)$, so this map is surjective. An $R = \mathbb{T}$ **theorem** says that this map identifies $\mathbb{T}_{\mathfrak{m}}$ with a certain quotient of $R_{\overline{\rho}}$ determined by local data. (In practice one needs some more flexibility, such as to include a Hecke operator at p, or to impose determinant conditions, to invert p before claiming to have an isomorphism, etc.)

Lecture 3: Galois deformation rings

Mok
October 16, 2009
Notes by Sam Lichtenstein

Let G be a profinite group and $\overline{\rho} : G \to \mathrm{GL}_n(k)$ a representation defined over a finite field k of characteristic p. Let Λ be a complete discrete valuation ring with residue field k, e.g. $\Lambda = W(k)$. Let $\mathcal{C}_\Lambda$ be the category of artinian local Λ-algebras with residue field k, and local morphisms. Let $\widehat{\mathcal{C}}_\Lambda$ be the category of complete Noetherian local Λ-algebras with residue field k, i.e. the pro-category of $\mathcal{C}_\Lambda$.

1. Deformation functors

Define $\mathrm{Def}(\overline{\rho}) : \widehat{\mathcal{C}}_\Lambda \to \mathrm{Sets}$ by

$$\mathrm{Def}(\overline{\rho})(A) = \{(\rho, M, \iota)\}/\sim$$

where M is a free A-module of rank n, $\rho : G \to \mathrm{GL}_A(M)$ is a continuous representation, $\iota : \rho \otimes_A k \cong \overline{\rho}$ is an isomorphism, and two such triples are equivalent when the representations are isomorphic in a manner which respects the ι's. Define the **framed** deformation functor $\mathrm{Def}^\square(\overline{\rho})$ by

$$\mathrm{Def}^\square(\overline{\rho})(A) = \{(\rho, M, \iota, \beta)/\sim$$

where β is a basis for M lifting the standard basis for k^n under ι. Morally, $\mathrm{Def}^\square$ is the set of liftings of $\overline{\rho}$ into $\mathrm{GL}_n(A)$.

There is a forgetful functor $\mathrm{Def}^\square \to \mathrm{Def}$.

Equivalent definitions are

$$\mathrm{Def}^\square(\overline{\rho})(A) = \{\rho : G \to \mathrm{GL}_n A \,|\, \rho \bmod \mathfrak{m}_A = \overline{\rho}\},$$

$$\mathrm{Def}(\overline{\rho})(A) = \mathrm{Def}^\square(\overline{\rho})(A)/(\text{conjugation by } \Gamma_n(A) := \ker(\mathrm{GL}_n(A) \to \mathrm{GL}_n(k))).$$

Note: it is easy to see that $\mathrm{Def}^\square(\overline{\rho})(A) = \varprojlim_i \mathrm{Def}^\square(\overline{\rho})(A/\mathfrak{m}_A^i)$. It is also true (but requires an argument) that $\mathrm{Def}(\overline{\rho})(A) = \varprojlim_i \mathrm{Def}(\overline{\rho})(A/\mathfrak{m}_A^i)$. In other words "we can compute these functors on the level of artinian quotients", so we just need to consider them on the category $\mathcal{C}_\Lambda$.

2. p-finiteness

We cannot hope to represent $\mathrm{Def}(\overline{\rho})$ or $\mathrm{Def}^\square(\overline{\rho})$ in $\widehat{\mathcal{C}}_\Lambda$ (which only contains Noetherian rings) unless G is "not too big".

Definition. We say G **satisfies the p-finiteness condition** if for every open subgroup $H \subset G$ of finite index, there are only finitely many continuous group homomorphisms $H \to \mathbb{Z}/p\mathbb{Z}$ (i.e., only finitely many open subgroups of index p). (This holds if and only if for any such H, the maximal pro-p quotient of H is topologically finitely generated.)

We are interested in two cases.

(1) $G = G_K$ for a local field K finite over $\mathbb{Q}_\ell$ (allowing $\ell = p$!).
(2) $G = G_{K,S}$ for a number field K and S a finite set of ramified primes.

In case (1), $H = G_{K'}$ for a finite extension K'/K, and the p-finiteness condition follows from the fact that the local field K' of characteristic 0 has only finitely many extensions of any given degree (such as degree p). For (2), H corresponds to some finite extension K'/K unramified outside of S, so the index-p open subgroups of H correspond to certain degree-p extensions of K' unramified away from the places of K' over S. Thus, the p-finiteness follows from the **Hermite-Minkowski theorem**, which says that only finitely many extensions of K of bounded degree unramified outside S.

Returning to the general situation, assume G satisfies p-finiteness. By *Schlessinger's criterion*, we will eventually see that $\mathrm{Def}^\square(\overline{\rho})$ is always representable in $\widehat{\mathcal{C}}_\Lambda$, so there exists a universal framed deformation ring $R_{\overline{\rho}}^\square \in \widehat{\mathcal{C}}_\Lambda$ and a universal framed deformation $\rho_{\overline{\rho}}^\square$ satisfying the natural universality property. We will also see that $\mathrm{Def}(\overline{\rho})$ is itself representable by a universal deformation ring $(R_{\overline{\rho}}, \rho_{\mathrm{univ}})$, at least when $\mathrm{End}_G(\overline{\rho}) = k$. This will be the case if $\overline{\rho}$ is absolutely irreducible, and also if $n = 2$ and $\overline{\rho}$ is a non-split extension of distinct characters.

3. Zariski tangent space to the deformation functors

Let $k[\epsilon]$ denote the ring of dual numbers of k. The **tangent space** to a functor $F : \widehat{\mathcal{C}}_\Lambda \to Sets$ is $F(k[\epsilon]) =: t_F$. Initially this is just a set; the hypotheses of Schlessinger's criterion give it a natural structure of k-vector space (compatibly with natural transformations in F).

Let $V \in \mathrm{Def}(\overline{\rho})(k[\epsilon]) = t_{\mathrm{Def}(\overline{\rho})}$. Then by definition there is given a specified isomorphism $V/\epsilon V \cong \overline{\rho}$, so we obtain an exact sequence

$$0 \to \epsilon V \to V \to \overline{\rho} \to 0.$$

But it is easy to see that ϵV is naturally $k[G]$-isomorphic to $\overline{\rho}$ as well. Hence we see

$$t_{\mathrm{Def}(\overline{\rho})} = \mathrm{Ext}^1_{k[G]}(\overline{\rho}, \overline{\rho}) = \mathrm{H}^1(G, \mathrm{Ad}(\overline{\rho}));$$

this respects the k-linear structure on both sides.

More explicitly, given $\rho \in \mathrm{Def}^\square(\overline{\rho})(k[\epsilon])$ we can write $\rho(g) = \overline{\rho}(g) + \epsilon\Phi(g)\overline{\rho}(g)$ for $\Phi(g) \in \mathrm{Ad}(\overline{\rho})$. One can compute that the condition that ρ is a group homomorphism is the 1-cocycle condition on Φ. So $t_{\mathrm{Def}^\square(\overline{\rho})} = Z^1(G, \mathrm{Ad}(\overline{\rho}))$. Similarly one checks that two framed deformations are conjugate under $\Gamma_0(k[\epsilon]) = I_n + \epsilon M_n(k)$ if and only if their associated cocycles differ by a 1-coboundary. We conclude that $t_{\mathrm{Def}(\overline{\rho})} = \mathrm{H}^1(G, \mathrm{Ad}(\overline{\rho}))$, and

$$\dim_k B^1(G, \mathrm{Ad}(\overline{\rho})) = \dim \mathrm{Ad}(\overline{\rho}) - \dim \mathrm{H}^0(G, \mathrm{Ad}(\overline{\rho}))$$

is the **number of framed variables**. The p-finiteness hypothesis says precisely that $\dim Z^1, \dim \mathrm{H}^1 < \infty$. If moreover $\mathrm{End}_G(\overline{\rho}) = k$ then $h^0(G, \mathrm{Ad}(\overline{\rho})) = 1$, and we are in the representable situation. The forgetful functor $\mathrm{Def}^\square(\overline{\rho}) \to \mathrm{Def}(\overline{\rho})$ induces a map $R_{\overline{\rho}} \to R^\square_{\overline{\rho}}$, which turns out to be formally smooth, and thus realizes $R^\square_{\overline{\rho}}$ as a ring of formal power series (in some number d of variables) over $R_{\overline{\rho}}$. The number d is precisely the number of framed variables, which in this case is $n^2 - 1$.

Concretely, what is going on is that if $\overline{\rho}$ has only scalar endomorphism (so likewise for any lifting of $\overline{\rho}$) and we consider the universal deformation $R_{\overline{\rho}}$ then to "universally" specify a basis which residually lifts the identity is precisely to applying conjugation by a residually trivial matrix which is unique up to a unit scaling factor. And we can eliminate the unit scaling ambiguity by demanding (as we always may in a unique way) that the upper left matrix entry is not merely a unit but is equal to 1. Thus, the framing amounts to specifying a "point" of the formal $R_{\overline{\rho}}$-group of PGL_n at the identity, which thereby proves the asserted description of the universal framed deformation ring in these cases as a formal power series ring over $R_{\overline{\rho}}$ in $n^2 - 1$ variables. To be explicit, over

$$R^\square(\overline{\rho}) = R(\overline{\rho})[\![Y_{i,j}]\!]_{1 \le i,j \le n, (i,j) \neq (1,1)}$$

the universal framed deformation is the lifting $\overline{\rho}_{\mathrm{univ}}$ equipped with the basis obtained from the standard one by applying the invertible matrix $1_n + (Y_{i,j})$ where $Y_{1,1} := 0$.

It must be stressed that we will later need to work with cases in which $\overline{\rho}$ is *trivial* (of dimension 2), so $R_{\overline{\rho}}$ does not generally exist. This is why the framed deformation ring is useful.

4. References

- Mazur's articles in "Galois groups over $\mathbb{Q}$" and "Modular Forms and Fermat's Last Theorem".
- Kisin's notes from CMI summer school in Hawaii.

5. More on Zariski tangent spaces to deformation functors

From now on fix G to be either G_K for local K or $G_{K,S}$ for a number field K. Fix $\overline{\rho} : G \to \mathrm{GL}_n(k)$ and suppose the characteristic of the finite field k is p. If F is a deformation functor represented by $R \in \widehat{\mathcal{C}}_\Lambda$, recall that

$$F(A) = \mathrm{Hom}_{\Lambda\text{-alg}}(R, A), \qquad t_F = F(k[\epsilon]) = \mathrm{Hom}_{\Lambda\text{-alg}}(R, k[\epsilon]) = \mathrm{Hom}_{\Lambda\text{-alg}}(R/(\mathfrak{m}_R^2 + \mathfrak{m}_\Lambda R), k[\epsilon]).$$

The last equality is because the Λ-algebra maps are *local* morphisms, so in particular they send $\mathfrak{m}_R$ to $\epsilon k[\epsilon]$, and hence $\mathfrak{m}_R^2$ to zero. But by general nonsense we have

$$R/(\mathfrak{m}_R^2 + \mathfrak{m}_\Lambda R) = k \oplus \frac{\mathfrak{m}_R}{\mathfrak{m}_R^2 + \mathfrak{m}_\Lambda R},$$

where the second summand is square zero. Thus we see

$$t_F = \mathrm{Hom}_k(\frac{\mathfrak{m}_R}{\mathfrak{m}_R^2 + \mathfrak{m}_\Lambda R}, k) = t_R^*,$$

where for $A \in \widehat{\mathcal{C}}_\Lambda$ we define the **reduced Zariski cotangent space of** A to be

$$t_A^* = \frac{\mathfrak{m}_A}{\mathfrak{m}_A^2 + \mathfrak{m}_\Lambda R}.$$

Exercise. Fix a map $A \xrightarrow{f} B$ in $\widehat{\mathcal{C}}_\Lambda$. Then f is surjective if and only if $t_f^* : t_A^* \to t_B^*$ is surjective. [Use completeness... it's a Nakayamal's lemma sort of thing.]

A corollary of the Exercise is that if $d = \dim_k t_F = \dim_k t_R^*$ then we can pick a k-basis $x_1, \ldots, x_d$ of t_R^*, lift it to a collection $\widetilde{x}_i \in \mathfrak{m}_R \subset R$, and then the map $\Lambda[\![X_1, \ldots, X_d]\!] \to R$ sending X_i to $\widetilde{x}_i$ will be surjective. *A priori* bounds for the number of generators in the kernel (and hence on the dimension of R) can be obtained by estimating certain H^2s in the cohomology of G, which will be discussed later. These dimension bounds are sometimes useful, but usually not strong enough to give good control on R.

6. Examples

A local case. Let $K/\mathbb{Q}_\ell$ be local with $\ell \neq p$ and $G = G_K$. Let $\overline{\rho}$ be the trivial representation of dimension n. Then in particular $\mathrm{End}_G \overline{\rho} \supsetneq k$, so only the framed deformation functor is representable. In this case we can actually construct $R^\square(\overline{\rho})$ by hand. If $\rho : G \to \mathrm{GL}_n A$ is a deformation of the trivial representation $\overline{\rho}$, then G lands in the kernel $\Gamma_n A \subset \mathrm{GL}_n A$. Now $\Gamma_n A = I_n + M_n(\mathfrak{m}_A)$, explicitly, which is a pro-p group isomorphism to the additive group $M_n(A)$. In particular ρ factors throug h the maximal pro-p quotient of G.

In particular $\rho|_{I_K}$ factors through the p-part of the tame quotient $I_K^{\mathrm{tame}} = I_K/I_K^{\mathrm{wild}}$ of the inertia I_K of K. The picture to keep in mind is the tower of field extensions

$$K \hookrightarrow K^{\mathrm{unr}} \hookrightarrow K^{\mathrm{tame}} \hookrightarrow \overline{K}.$$

Now from the structure of local fields we know that the p-part of I_K^{tame} is

$$I_K^{\mathrm{tame},(p)} = \mathbb{Z}_p(1).$$

Here the twist means that if $\sigma \in I_K^{\mathrm{tame},(p)}$ then $\mathrm{Frob}_K \, \sigma \, \mathrm{Frob}_K^{-1} = \sigma^q$ where $q = \ell^r = \#(\mathcal{O}_K/\mathfrak{m}_K)$. Fix a lift $f \in G$ of Frob_K and τ a topological generator of $I_K^{\mathrm{tame},(p)}$. What we can conclude is that a lift ρ to any A is specified by the images of f and τ, subject to the relation

$$\rho(f)\rho(\tau) = \rho(\tau)^q \rho(f).$$

So we can take

$$R^\square(\overline{\rho}) = \Lambda[\![\{f_{ij}, \tau_{ij}\}_{1 \leq i,j \leq n}]\!]/I$$

where the ideal of relations I is generated by the ones given by the matrix equations

$$[I_n + (f_{ij})][I_n + (\tau_{ij})] = [I_n + (\tau_{ij})]^q [I_n + (f_{ij})].$$

A global case. For a global case we'll consider *characters* of $G = G_{K,S}$. Note that we have a wonderful fact in this case. The Teichmüller lift $[\cdot] : k \to W(k)$ is a multiplicative section of $W(k) \to k$. This allows us to twist any character $\overline{\rho}$ by the Teichmüller lift $[\overline{\rho}^{-1}]$ of its reciprocal. to conclude that $R(\overline{\rho}) = R(\mathbf{1})$ where $\mathbf{1} : G \to k^\times$ is the trivial character. In other words, the universal deformation of a character $\overline{\rho}$ is just a twist of the universal deformation of the trivial character (using the same coefficient ring).

Arguing just like in the local case, it follows that any lift ρ to A of the trivial mod p character $\overline{\rho}$, must factor through the maximal pro-p quotient $G_{K,S}^{\mathrm{ab},(p)}$ of the abelianization of $G_{K,S}$.

Let us specialize now to the case $K = \mathbb{Q}$ [the case of a general number field is similar, but requires class field theory]. Assume $p \in S$, since otherwise stuff is boring. By the Kronecker-Weber theorem we know that $G_{\mathbb{Q},S}^{\mathrm{ab}} = \prod_{\ell \in S} \mathbb{Z}_\ell^\times$, which implies that the maximal pro-p quotient is

$$G_{\mathbb{Q},S}^{\mathrm{ab},(p)} = \prod_{\ell \in S, \ell \equiv 1 (p)} (\mathbb{F}_\ell^\times)^{(p)} \times (1 + p\mathbb{Z}_p).$$

So we can, in this case, simply take $R = \Lambda[\![G_{\mathbb{Q},S}^{\mathrm{ab},(p)}]\!]$ to be the formal group algebra over Λ. From the description of $G_{\mathbb{Q},S}^{\mathrm{ab},(p)}$ we can be very explicit:

$$R = \frac{\Lambda[\![\{X_\ell\}_{\ell\in S,\ell\equiv 1(p)}, T]\!]}{(\{((X_\ell+1)^{p^{\mathrm{ord}_p(\ell-1)}}-1)\}_{\ell\in S,\ell\equiv 1(p)})}$$

In particular if $S = \{p,\infty\}$ then $R \cong \Lambda[\![T]\!]$.

For a general number field K this relates to the *Leopoldt conjecture* which says that $\mathrm{rk}_{\mathbb{Z}_p}(G_{K,S}^{\mathrm{ab},(p)}) = 1+r_2$, where r_2 is the number of conjugate pairs of complex embeddings of K.

7. Local and Global

We can relate the two examples from the last subsection in the following manner, which will be extremely important later in one of Kisin's key improvements of Wiles' method. Let $G = G_{K,S}$, $\overline{\rho} : G_{K,S} \to \mathrm{GL}_n(k)$ a fixed residual representations, and Σ a finite set of primes. For each $v \in \Sigma$ we have

$$\overline{\rho}|_{G_v} : G_{K_v} = G_v \hookrightarrow G_K \twoheadrightarrow G_{K,S} \xrightarrow{\overline{\rho}} \mathrm{GL}_n(k).$$

We have local framed deformation rings $R_v^{\square} := R^{\square}(\overline{\rho}|_{G_v})$. Define a variation of the global framed deformation functor by

$$\mathrm{Def}^{\square,\Sigma}(\overline{\rho})(A) = \{(\rho_A, \{\beta_v\}_{v\in\Sigma})\}/\sim;$$

here, ρ_A is a deformation of $\overline{\rho}$ and β_v is a basis for $\rho_A|_{G_v}$ which reduces to the standard basis for $\overline{\rho}$. Then in fact $\mathrm{Def}^{\square,\Sigma}(\overline{\rho})$ is also representable, by a ring $R_{K,S}^{\square,\Sigma}$. For each $v \in \Sigma$ we have a forgetful map

$$\mathrm{Def}^{\square,\Sigma}(\overline{\rho}) \to \mathrm{Def}(\overline{\rho}|_{G_v})$$

and hence on the revel of representing objects, an algebra

$$R_v^{\square} \to R_{K,S}^{\square,\Sigma}.$$

In concrete terms, this is saying that if we form the universal deformation of $\overline{\rho}$ equipped with a framing along Σ and then forget the framing away from v and restrict to G_v, the resulting framed deformation of $\overline{\rho}|_{G_v}$ with coefficients in $R_{K,S}^{\square,\Sigma}$ is uniquely obtained by specializing the universal framed deformation of $\overline{\rho}|_{G_v}$ along a unique local Λ-algebra homomorphism $R_v^{\square} \to R_{K,S}^{\square,\Sigma}$.

Hence, by the universal property of completed tensor products (to be discussed in Samit's talk rather generally) we get an important map

$$\widehat{\bigotimes}_{\Lambda} R_v^{\square} \to R_{K,S}^{\square,\Sigma}$$

in $\widehat{\mathcal{C}}_\Lambda$. (Note that we have to take the completion of the algebraic tensor product, which is not itself a complete ring. For example, $\Lambda[\![x]\!] \otimes_\Lambda \Lambda[\![y]\!]$ is a gigantic non-noetherian ring, but the corresponding completed tensor product is $\Lambda[\![x,y]\!]$.) This is a rather interesting extra algebra structure on the global framed deformation ring, much richer than its mere Λ-algebra structure; of course, this all has perfectly good analogues without the framings, assuming that $\overline{\rho}$ and its local restrictions at each $v \in \Sigma$ have only scalar endomorphisms.

This idea of viewing a global deformation ring as an algebra over a (completed) tensor product of local deformation rings is the key to Kisin's method for "patching" deformation rings in settings going far beyond the original Taylor-Wiles method (where only the Λ-algebra structure was used).

Lecture 4: Generic fibers of deformation rings

Brian
October 23, 2009
Notes by Sam Lichtenstein

1. Some observations

Fix $\overline{\rho} : G_{\mathbb{Q},S} \to \mathrm{GL}_2(k)$ absolutely irreducible, and let $\rho : G_{\mathbb{Q},S} \to \mathrm{GL}_2(R)$ be the universal deformation. We're interested in the map $R \to \mathbb{T}_{\mathfrak{m}}$ for some Hecke algebra defined in terms of $\overline{\rho}$. Note that the Hecke algebra is 1-dimensional, and even finite free over $\mathbb{Z}_p$. The universal deformation ring R, however, often has dimension > 1 and nonzero p-torsion. In other words, the surjection $R \twoheadrightarrow \mathbb{T}_{\mathfrak{m}}$ is not even close to being an isomorphism in general.

Example. Consider $X_0(49)$ which is an elliptic curve. [Cf. Nigel Boston's papers on explicit deformation rings for the details of this example.] Let $\overline{\rho}$ the representation from the 3-torsion of E, and let $S = \{3, 7, \infty\}$. Boston computed the universal deformation as

$$\rho : G_{\mathbb{Q},S} \to \mathrm{GL}_2\left(\frac{\mathbb{Z}_3[\![x_1, x_2, x_3]\!]}{(1+x_1)^3 - 1}\right).$$

Just by looking at the ring on the right side, it's clear that its dimension is at least 2. (This example doesn't illustrate the phenomenon of p-torsion, but oh well...)

Morally, the reason for the higher dimension of R is that we are not imposing any local conditions at all for the places in S.

A key observation is that even when we succeed in proving a modularity lifting theorem, we don't know until we're done that R is $\mathbb{Z}_p$-finite and flat. In other words, even when in fact R turns out to be nice, we have very little grasp of why it is nice without proving an $R = \mathbb{T}$ theorem.

However, this is really not so bad. For example, if we could show that $R[1/p] \cong \mathbb{T}_m[1/p]$, that's totally fine. After all, we're trying to study deformations of $\overline{\rho}$ over p-adic integer rings, which are p-torsion free and reduced, so we rig the Hecke algebra to have the same properties. In other words, **we only care about the "p-adic points" of** R so we can just as well study the structure of $R[1/p]$/nilpotents. And via Kisin's methods, it turns out that a thorough understanding of the "structure" of this ring is attainable in interesting cases and is exactly what is needed for modularity lifting theorems. Things we would like to know:

- Characterize in some moduli-theoretic manner the connected components of its spectrum (e.g., so we can detect when two p-adic points lie on the same component).
- Dimension.
- Singularities, i.e. the extent to which an appropriately defined notion of smoothness fails to hold.

For the last point, it is just as good in practice to pass to a formally smooth R-algebra (such as a power series ring over R). So we can consider the framed deformation ring.

Remark. A key point is that $R[1/p]$ is **very** far from being a local ring. For example, say $R = \mathbb{Z}_p[\![x]\!]$ (which is a rough prototype of the sort of ring that arises). Then

$$R[1/p] = \mathbb{Z}_p[\![x]\!][1/p] = \{f \in \mathbb{Q}_p[\![x]\!] \mid \text{denominators are } \textit{bounded} \text{ powers of } p\} \subsetneq \mathbb{Q}_p[\![x]\!].$$

This ring has lots of $\mathbb{Q}_p$-algebra maps $\mathbb{Z}_p[\![x]\!][1/p] \twoheadrightarrow \mathcal{O}_K[1/p]$ for finite extensions $K/\mathbb{Q}_p$, sending x into $\mathfrak{m}_K$. Hence it has lots of maximal ideals.

2. Digression on Jacobson rings

Definition. A **Jacobson ring** is a Noetherian ring A such that any $\mathfrak{p} \in \operatorname{Spec} A$ is the intersection of the maximal ideals containing $\mathfrak{p}$.

Clearly a quotient of a Jacobson ring is Jacobson. Less evident, but in the exercises of Atiyah-MacDonald, is that a finitely generated algebra over a Jacobson ring is Jacobson. Note that any field is Jacobson, as is any Dedekind domain with *infinitely many* primes (but not a dvr, nor a local ring which is not 0-dimensional!). In particular, a general localization of a Jacobson ring is certainly not Jacobson, though localization at a single element is (since it is a finitely generated algebra).

A consequence of the definition is that if $X_0 = \mathrm{MaxSpec}(A) \overset{j}{\hookrightarrow} \mathrm{Spec}\, A = X$, then j is a dense *quasihomeomorphism*, which means that $U_0 = X_0 \cap U \leftrightarrow U$ is a bijection between the collections of open sets in X_0 and X. Jacobson rings abstract the nice properties enjoyed by algebras of finite type over a field.

Claim. If R is a quotient of a formal power series ring over a complete dvr A with uniformizer π then $R[1/\pi]$ is Jacobson, and for all maximal ideals $\mathfrak{m} \subset R[1/\pi]$, the quotient $R[1/\pi]/\mathfrak{m}$ is finite over the fraction field $K = A[1/\pi]$ of A. Moreover, every K-algebra map from $R[1/\pi]$ to a finite extension K' of K carries R into the valuation ring A' of K', with the map $R \to A'$ actually a local map.

Note it is elementary that every K-algebra map from $R[1/\pi]$ to a finite extension K' of K has kernel that is maximal: the kernel P is at least a prime ideal, and $R[1/\pi]/P$ is an intermediate ring between the field K and the field K' of finite degree over K, so it is a domain of finite dimension over a field (namely K) and hence is itself a field. Hence, P is maximal.

Also, everything in the Claim can be deduced from facts in rigid geometry concerning K-affinoid algebras, by using the approach in deJong's IHES paper *Crystalline Dieudonné theory via formal and rigid geometry*. For convenience, we give a direct proof using commutative algebra, avoiding rigid geometry (but inspired by it for some of the arguments).

The proof of the Claim is somewhat long (and was omitted in the lecture).

Proof. To prove the claim, first note that if the claim holds for R then it holds for any quotient of R. Hence, it suffices to treat the case when $R = A[\![x_1, \ldots, x_n]\!]$ is a formal power series ring over A. We first check the more concrete second part of the Claim: for finite K'/K, any K-algebra map $R[1/\pi] \to K'$ carries R into the valuation ring A' of K' with $R \to A'$ moreover a local map. In other words, we are studying A-algebra maps $R \to K'$. This can be uniquely "promoted" to an A'-algebra map

$$A' \otimes_A R \to K',$$

and we can pass the tensor product through the "formal power series" formation since A' is a finite free A-module. In other words, we can rename A' as A to reduce to the case $K' = K$. So we claim that any A-algebra map $R \to K$ must be "evaluation" at an n-tuple in the maximal ideal of A. If we can show it carries each x_i to some a_i in the maximal ideal of A then the map kills $x_i - a_i$ for all i. By completeness of R it would be legal to make a "change of variables" renaming $x_i - a_i$ as x_i to reduce to the case when the map kills all x_i's. Since the quotient of R by the ideal generated by the x_i's is identified with A, after inverting π we get K (as a K-algebra!), so we'd have proved what we want.

Let's now show that indeed each x_i is carried to some a_i in the maximal ideal of A. By composing the given A-algebra map $R \to K$ with the natural inclusion $A[\![x_i]\!] \to R$ we are reduced to the case $n = 1$. That is, we wish to prove that any A-algebra map $A[\![x]\!] \to K$ must carry x to an element a in the maximal ideal of A. This map must kill some nonzero $f \in A[\![x]\!]$, as $A[\![x]\!][1/\pi]$ has infinite K-dimension as a vector space, and we can write $f = \pi^e f_0$ for some $e \geq 0$ and some f_0 not divisible by π. Thus, f_0 also dies in K, so by renaming it as f we arrange that f has some coefficient not divisible by π. This coefficient must occur in positive degree, as otherwise f would be a unit, which is absurd (as it is in the kernel of a map to a field). Now by the formal Weierstrass Preparation Theorem (in one variable –see Lang's *Algebra*), if $d > 0$ is the least degree of a coefficient of f not divisible by π then f is a unit multiple of a "distinguished" polynomial: a monic polynomial in x of degree d over A with all lower-degree coefficients divisible by π. Scaling away the unit, we can assume that f is a monic polynomial of degree $d > 0$ with all lower-degree coefficients divisible by π. Hence, $A[\![x]\!]/(f) = A[x]/(f)$ by long-division of formal power series (thanks to completeness of A!). Our map of interest therefore "is" an A-algebra map

$$A[x]/(f) \to K$$

and so it carries x to an element a of K that is a root of f. Since f is monic over A, we see $a \in A$. Since f has all lower-degree coefficients in the maximal ideal, necessarily a is in the maximal ideal too. That completes the proof of the second part of the Claim.

Now it remains to show the first part of the Claim: R is Jacobson, and if M is a maximal ideal of $R[1/\pi]$ then $R[1/\pi]/M$ is of finite degree over $A[1/\pi] = K$. We argue by induction on the number n of variables (motivated by the method of proof of the analytic Weierstrass Preparation theorem over $\mathbb{C}$ or non-archimedean fields), the case $n = 0$ being trivial. Also, it is harmless (even for the Jacobson property)

to make a finite extension on K if we wish. We will use this later, to deal with a technical problem when the residue field k is finite (which is of course the case of most interest to us).

Assume $n > 0$, and consider a nonzero $f \in R = A[\![x_1, \ldots, x_n]\!]$ contained in some chosen nonzero prime or maximal ideal; clearly f can be scaled by π-powers so it is not divisible by π. We want to get to the situation in which f involves a monomial term that is just a power of a single variable. Pick a monomial of least total degree appearing in f with coefficient in $A^\times$. (Such a term exists, since f is not divisible by π.) This least total degree d must be positive (as otherwise $f(0) \in A^\times$, so $f \in R^\times$, a contradiction). By relabeling, we may suppose x_1 appears in this monomial. If $n = 1$, this term is an $A^\times$-multiple of a power of x_1, so we're happy. Now assume $n > 1$ and consider the *homogeneous* change of variables which replaces x_i with $x_i + c_i x_1$ for all $i > 1$ (and leaves x_1 alone), with $c_i \in A$ to be determined in a moment. Each degree-d monomial

$$a_I x_1^{i_1} \cdots x_n^{i_n}$$

in f (before the change of variable) with total degree d contributes

$$a_I c_2^{i_2} \cdots c_n^{i_n} x_1^d$$

to the x_1^d term after the change of variable (with $i_1 = d - (i_2 + \cdots + i_n)$). All other monomials can only contribute to x_1^d with coefficient in maximal ideal of A. Thus, these other terms can be ignored for the purpose of seeing if we get x_1^d to appear with an $A^\times$-coefficient after the change of variables.

To summarize (when $n > 1$), whatever c_i's we choose in A, we get after change of variable that x_1^d appears with coefficient $h(c)$ for some polynomial h in $n-1$ variables over A that has *some* coefficient in $A^\times$ (since i_1 is determined by $i_2, \ldots, i_n$). Thus, h has nonzero reduction as a polynomial over the residue field k of A, so as long as this reduction is nonzero at some point in k^{n-1} we can choose the c's to lift that into A^{n-1} to get the coefficient of x_1^d to be in $A^\times$. If k is infinite, no problem. If k is finite (case of most interest!), for some finite extension k' of k we can find the required point in k'^{n-1}, so go back and replace A with the corresponding unramified extension (and the chosen prime with each of the ones over it after scalar extension) to do the job.

The upshot is that after a suitable change of variables (and possible replacement of A with a finite extension in case k is finite), we can assume that f contains some x_1^d with an $A^\times$-coefficient. Thus, if we view f in

$$R = (A[\![x_2, \ldots, x_n]\!])[\![x_1]\!]$$

then it satisfies the hypotheses of the general Weierstrass Preparation (with complete coefficient ring) as in Lang's *Algebra*. This implies that f is a unit multiple of a *monic* polynomial in x_1 whose lower-degree coefficients are in the maximal ideal of $R' = A[\![x_2, \ldots, x_n]\!]$ (which means A if $n = 1$). We can therefore scale away the unit so that f is such a "distinguished" polynomial, and then do long division in $R'[\![x_1]\!]$ due to completeness of R' to infer that

$$R/(f) = R'[\![x_1]\!]/(f) = R'[x_1]/(f).$$

This is a finite free R'-module!

We may now draw two consequences. First, if P is a prime ideal of $R[1/\pi]$ containing f then $R[1/\pi]/P$ is module-finite over the ring $R'[1/\pi]$ which is Jacobson by induction, so $R[1/\pi]/P$ is Jacobson. Hence, P is the intersection of all maximals over it, whence we have proved that $R[1/\pi]$ is Jacobson. Second, for a maximal ideal M of $R[1/\pi]$ containing f, the ring map $R'[1/\pi] \to R[1/\pi]/M$ is module-finite so its prime ideal kernel is actually maximal. That is, we get a *maximal* ideal M' of $R'[1/\pi]$ such that $R'[1/\pi]/M' \to R[1/\pi]/M$ is of finite degree. By induction, $R'[1/\pi]/M'$ is of finite degree over K, so we are done. □

3. Visualizing $R[1/p]$

Let $R = A[\![x_1, \ldots, x_n]\!]/(f_1, \ldots, f_m)$ and K be in the last subsection. Observe that $\mathrm{Hom}_{\mathrm{loc}.A-\mathrm{alg}}(R, A') = \mathrm{Hom}_{\mathrm{Frac}(A)-\mathrm{alg}}(R[1/\pi], A'[1/\pi] = \mathrm{Frac}(A'))$ for any finite dvr extension A' of A. This suggests the following geometric perspective on the ring $R[1/\pi]$: it corresponds to the locus of geometric points (x_i) with coordinates in $\overline{\mathrm{Frac}(A)}$ lying in the open polydisk $\{|x_1|, \ldots, |x_n| < 1\}$ at which the convergent power series $f_1, \ldots, f_m$ all vanish. To make this viewpoint precise, one must regard the spaces in question as **rigid analytic spaces**.

4. Final thought

We'll see that for Galois deformation rings R, the completions of $R[1/p]$ at maximal ideals are deformation rings for *characteristic zero* representations corresponding to the maximal ideals in question. This is very interesting, since R itself was entirely about deforming $\bmod p$ things!

5. Back to examples of explicit universal deformation rings

Caveat: These sorts of examples are kind of "useless". The reference for N. Boston's examples is *Inv. Math.* **103** (1991).

Example 1 [loc. cit., Prop. 8.1.] Let $E : y^2 = x(x^2-8x+8)$, an elliptic curve with complex multiplication by $\mathbb{Q}(\sqrt{-2})$. Let $\overline{\rho}$ be the representation on the 3-torsion:

$$G_{\mathbb{Q},\{2?,3,5,\infty\}} \to \mathrm{GL}_2(\mathbb{F}_3).$$

In general we know that there is some surjection $\mathbb{Z}_3[\![T_1,\ldots,T_d]\!] \twoheadrightarrow R(\overline{\rho})$ where we know the smallest d is (by NAK) $d = \dim \mathfrak{m}_R/(\mathfrak{m}_R^2, 3)$, and $\mathfrak{m}_R/(\mathfrak{m}_R^2, 3) = \mathrm{H}^1(G_{\mathbb{Q},\{2?,3,5,\infty\}}, \mathrm{Ad}(\overline{\rho}))$. Here the adjoint module is $\mathrm{Ad}(\overline{\rho}) = \mathrm{End}_{\mathbb{F}_3}(\overline{\rho})$ with $G_{\mathbb{Q},\{2?,3,5,\infty\}}$ acting by conjugation via $\overline{\rho}$. In this particular case one can compute that $d = 5$, so

$$R(\overline{\rho}) = \mathbb{Z}_3[\![T_1,\ldots,T_5]\!]/I$$

where the ideal of relations has the form

$$I = \delta \cdot (f, g)$$

for

$$f = 8u^4 - 8u^2 + 1, g = 8e^3 - 4u, \qquad u = (1 + T_4T_5)^{1/2}$$

and δ (which may involve all the T_is) is obtained by choosing a certain presentation of a pro-3 group (coming from a wild inertia group, perhaps for the splitting field of $\overline{\rho}$?), and setting $\delta = \det(\rho^{\mathrm{univ}}(y) - 1)$ where y is a particular generator in said presentation. Consequently one can write down some "explicit" deformations of $\overline{\rho}$ by looking for solutions to the relations above in a $\mathbb{Z}_3$-algebra...

Example 2 [Boston-Ullom]. Let $E = X_0(49)$ and $\overline{\rho} = \overline{\rho}_{E,3}$ the representation on the 3-torsion:

$$G_{\mathbb{Q},\{3,7,\infty\}} \to \mathrm{GL}_2(\mathbb{F}_3).$$

In this case the universal deformation ring is particularly simple:

$$R \cong \mathbb{Z}_3[\![T_1,\ldots,T_4]\!]/((1+T_4)^3 - 1).$$

We have $(1+T_4)^3 - 1 = T_4(T_4^2+3T_4+3)$. The quadratic factor is irreducible over $\mathbb{Q}_3$, but not over $\mathbb{Q}_3(\sqrt{-3})$. So, loosely speaking, $\mathrm{Spec}\, R$ has two irreducible components but three "geometric" irreducible components: $T_4 = 0$ and T_4 equal to either of the conjugate roots of the quadratic factor. For example, to recover the 3-adic Tate module of E one considers the map $R \to \mathbb{Z}_3$ given by mapping all T_is to 0. This is a sort of "canonical" $\mathbb{Z}_3$-point of $\mathrm{Spec}\, R$. Since the quadratic factor of the relation is $\mathbb{Q}_3$-irreducible, so that quadratic field cannot be $\mathbb{Q}_3$-embedded into $\mathbb{Q}_3$, *every* $\mathbb{Z}_3$-point lies in the $T_4 = 0$ component.

The lesson to take from this seems to be that it can be hard to detect components, or more generally aspects of the geometry, of $\mathrm{Spec}\, R$, when only looking at p-adic points over a small field like $\mathbb{Q}_p$; we have to expect to work with points in many finite extensions in order to effectively probe the geometry. All this is by way of motivation for our interest in characteristic zero points of deformation rings, and (for example) our willingness to throw out all possible nastiness at p by studying $R[1/p]$ instead of R itself.

6. Back to Characteristic 0

Now let Λ be a p-adic dvr with fraction field K and residue field k. Let $R = \Lambda[\![X_n,\ldots,X_n]\!]/I$ be the universal deformation ring of a residual representation $\overline{\rho} : \Gamma \to \mathrm{GL}_N(k)$, for a profinite group Γ satisfying the requisite p-finiteness conditions (e.g. G_K for local K or $G_{K,S}$ for a number field K).

Remark. We have seen above that for any maximal ideal $\mathfrak{m} \subset R[1/p]$, the residue field $R[1/p]/\mathfrak{m}$ is of finite degree over k. The intuition for this fact is that these closed points of $\operatorname{Spec} R[1/p]$ correspond to Galois orbits over K of $\overline{K}$-solutions to $I = 0$ in the open unit n-polydisk. (The case $n = 1$ is a consequence of the Weierstrass Preparation Lemma. One *can* relate the geometry of $\operatorname{Spec} R[1/p]$ to the geometry of the aforementioned "rigid analytic space" $I = 0$. For example, if $R[1/p]$ is connected (no nontrivial idempotents) then $I = 0$ is connected in the sense of rigid geometry. One can also match up the dimensions of the components. The input for this equivalence is the (self-contained!) §7 of de Jong's IHES paper *Crystalline Dieudonné theory* ..., but we won't use it.

We also saw above that any K-algebra map $R[1/p] \to K'$ for a finite extension K'/K is actually given by sending all the X_is to elements $x_i \in \mathfrak{m}_{K'} \subset \mathcal{O}_{K'} \subset K'$. In other words, $R \subset R[1/p]$ actually lands in $\mathcal{O}_{K'}$!

Now fix a K-algebra map $x : R[1/p] \twoheadrightarrow K'$ into a finite extension of K. ("Contemplate a p-adic point of $\operatorname{Spec} R$".) Let

$$\rho_x : \Gamma \overset{\rho^{\text{univ}}}{\to} \mathrm{GL}_N(R) \to \mathrm{GL}_N(R[1/p]) \to \mathrm{GL}_N(K')$$

be the specialized representation. (In the Boston-Ullom example above, when we take $x : R[1/3] \to \mathbb{Q}_3$ to be the map sending all the T_is to zero, then ρ_x is the 3-adic Tate module of $X_0(49)$.)

Goal: Understand the dimension $\dim R[1/p]_{\mathfrak{m}_x} = \dim R[1/p]^\wedge_{\mathfrak{m}_x}$**.** (Here $(\cdot)^\wedge$ denotes completion.)

For instance, is this complete local ring regular? Perhaps even a power series ring over K'? If so, then its dimension is $\dim \mathfrak{m}_x/\mathfrak{m}_x^2$.

Theorem. *Let* $\rho_x^{\text{univ}} : \Gamma \to \mathrm{GL}_N(R[1/p]^\wedge_{\mathfrak{m}_x})$ *be induced from* ρ^{univ} *by the natural map* $R \to R[1/p]^\wedge_{\mathfrak{m}_x}$*. Then the diagram*

$$\begin{array}{ccc} \Gamma & \xrightarrow{\rho_x^{\text{univ}}} & \mathrm{GL}_N(R[1/p]^\wedge_{\mathfrak{m}_x}) \\ & \overset{\text{cont}}{\searrow} & \downarrow{\scriptstyle \rho_x} \\ & & \mathrm{GL}_N(K') \end{array}$$

commutes, and in fact ρ_x^{univ} *is the universal for continuous deformations of* ρ_x*.*

More precisely, if one considers the category $\widehat{\mathcal{C}}$ *of complete local noetherian* K'*-algebras with residue field* K'*, and the functor on the category* $\mathcal{C}$ *of artinian quotients of objects in* $\widehat{\mathcal{C}}$ *which picks out those deformations of* ρ_x *which are continuous for the* p*-adic topology on such artinian quotients, regarded as finite-dimensional* K'*-vector spaces, then* $R[1/p]^\wedge_{\mathfrak{m}_x}$ *is the representing object.*

Remark. If A is a complete local Noetherian F-algebra and the characteristic of F is zero, and $A/\mathfrak{m} = F'$ is a finite extension of F, then there exists a unique F-algebra lift $F' \hookrightarrow A$. Why? By completeness we have Hensel's lemma and by characteristic zero we have F'/F separable. So we can find solutions in A to the defining polynomial of F' over F.

Why do we care about the theorem?

(1) The deformation ring $R[1/p]^\wedge_{\mathfrak{m}_x}$ is isomorphic to $K'[\![T_1, \ldots, T_n]\!]$ if and only if $R[1/p]^\wedge_{\mathfrak{m}_x}$ is regular (by the Cohen structure theorem), and the power series description is precisely the condition that the corresponding deformation functor for ρ_x is formally smooth (i.e., no obstruction to lifting artinian points in characteristic 0). This holds precisely when $\mathrm{H}^2(\Gamma, \mathrm{Ad}(\rho_x)) = 0$. So that is interesting: a computation in Galois cohomology in characteristic 0 can tell us information about the structure of $R[1/p]$ at closed points.

(2) $(\mathfrak{m}_x/\mathfrak{m}_x^2)^\vee \cong \mathrm{H}^1_{\text{cont}}(\Gamma, \mathrm{Ad}(\rho_x))$, by the continuity condition we imposed on the deformations in the theorem.

Combining (1) and (2), we can check regularity of $R[1/p]$ at a closed point and in such cases then even compute $\dim_x R[1/p]$ by doing computations in (continuous) Galois cohomology with p-adic coefficients!

7. Proof of theorem

Step 1: Reduce to the case $K' = K$**.** Here is the trick. Set $\Lambda' = \mathcal{O}_{K'}$. Note that $\Lambda' \otimes_\Lambda R$ is local because $(\Lambda' \otimes_\Lambda R)/\mathfrak{m}_R = \Lambda' \otimes_\Lambda k = k'$ is a field. The Λ'-algebra $\Lambda' \otimes_\Lambda R$ is the universal deformation ring of $\overline{\rho} \otimes_k k'$

(where k' is the residue field of K') when using Λ'-coefficients; this behavior of deformation ring with respect to finite extension of the coefficients will be proved in Samit's talk. Consider the diagram

$$\begin{array}{ccc} K' \otimes_K R[1/p] & \xrightarrow{x'} & K' \\ \| & & \uparrow \\ (\Lambda' \otimes_\Lambda R)[1/p] & & \\ \uparrow & & \\ \Lambda' \otimes_\Lambda R & \longrightarrow & \Lambda' \end{array}$$

Exercise: $(\Lambda' \otimes_\Lambda R)[1/p]^\wedge_{\mathfrak{m}_{x'}} \cong R[1/p]^\wedge_{\mathfrak{m}_x}$ as K'-algebras.

So we can rename Λ' as Λ, completing the reduction.

Step 2: Observe that since $\overline{\rho}$ is absolutely irreducible, so is ρ_x. Consequently any deformation of ρ_x has only scalar endomorphisms.

Step 3: Consider any deformation

$$\begin{array}{ccc} \Gamma & \xrightarrow{\theta} & \mathrm{GL}_N(A) \\ & \searrow^{\rho_x} & \downarrow \\ & & \mathrm{GL}_N(K) \end{array}$$

where A is a finite local K-algebra with residue field K. We would like to show that there exists a unique K-map $R[1/p]^\wedge_{\mathfrak{m}_x} \to A$ which takes ρ_x^{univ} to θ, up to conjugation. Why is this sufficient? Because if so, then there would be lifts of ρ_x to $\mathrm{GL}_N(A)$, one coming from ρ_x^{univ} and the other being θ, which are $\mathrm{GL}_N(A)$-conjugate to one another by some matrix M. Upon reduction to $\mathrm{GL}_N(K)$, the matrix $\overline{M}$ would centralize ρ_x. So by Step 2, $\overline{M}$ must be a scalar endomorphism $c \in K^\times$. Consequently we can replace M by $c^{-1}M$ to conclude that the two lifts are conjugate to one another by a matrix which is residually trivial. The latter is precisely what we need to prove that ρ_x^{univ} is universal. (Note that if we used framed deformations throughout then this little step wouldn't be needed. It is important because in later applications we will certainly want to apply the Theorem to cases for which $\overline{\rho}$ is not absolutely irreducible. The reader can check that the proof of the Theorem works in the framed setting once the preceding little step is bypassed.)

The map we need is the same as making a local K-algebra map

$$R[1/p]_{\mathfrak{m}_x} \to A$$

with the same property with respect to θ, since A is a complete K-algebra. (Note that this "uncompletion" step is only possible since we already did Step 1! We originally completed $R[1/p]_{\mathfrak{m}_x}$, which is a K-algebra and generally not a K'-algebra.) The latter is the same as a K-algebra map $R[1/p] \to A$ such that $R[1/p] \to A \to K$ is the original point x, which takes ρ^{univ} to θ. ("It's all a game in trying to get back to R".) In other words, we wanted a dotted map in the diagram

$$\begin{array}{ccccc} R & \longrightarrow & R[1/p] & \overset{\exists!?}{\dashrightarrow} & A \\ {\scriptstyle \exists!\widetilde{x}}\downarrow & & & \searrow^{x} & \downarrow \\ \Lambda & & \longrightarrow & & K \end{array}$$

(The existence of $\widetilde{x}$ is by one of the propositions from §6.) But $R[1/p]$ is just a localization of R and A is a $\Lambda[1/p]$-algebra (it is a K-algebra!), so in fact the existence of a unique dotted map above is equivalent to the existence of a unique dotted map α in the diagram

$$\begin{array}{ccc} R & \overset{\exists!?}{\overset{\alpha}{\dashrightarrow}} & A \\ {\scriptstyle \widetilde{x}}\downarrow & & \downarrow \\ \Lambda & \longrightarrow & K \end{array}$$

such that α takes ρ^{univ} to θ. Now unfortunately A is not in the category $\widehat{\mathcal{C}}_\Lambda$ [typically it is something like $K[t]/(t^7)$], so θ is not quite a deformation of $\overline{\rho}$, so we cannot appeal directly to the universal property of (R, ρ^{univ}). Instead we need to mess around a bit.

Here's the point. $A = K \oplus \mathfrak{m}_A$ and $\mathfrak{m}_A$ is a finite-dimensional K-vector space which is nilpotent.

Claim. $\mathfrak{m}_A = \varinjlim I$ where the limit is taken over Λ-finite multiplicatively stable Λ-modules I.

(Idea of the proof: take products and products and more products. By nilpotence and finite-dimensionality of $\mathfrak{m}_A$ over K, you don't have to keep going forever. Then take the Λ-span of finite collections of such products to get the desired I's.)

Write Λ_I for $\Lambda \oplus I$.

Lemma/Exercise: Any Λ-algebra map $R \to A$ lands in some Λ_I. (Hint: choose I containing the images of all the X's.)

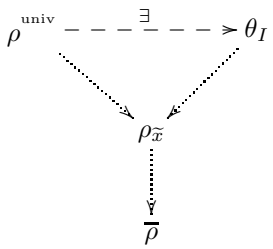
So it's enough to show two things.

(1) For some I we have a map $R \to \Lambda_I$ giving a deformation θ_I of the "integral lattice" version $\rho_{\widetilde{x}}$ of ρ_x. The image of Γ under ρ^{univ} is topologically finitely generated (since $\mathrm{GL}_N(R)$ is essentially pro-p and Γ satisfies the p-finiteness condition), so then there exists *some* I_0 such that θ factors through $\mathrm{GL}_N(\Lambda_{I_0})$, giving a map $\theta_{I_0} : \Gamma \to \mathrm{GL}_N(\Lambda_{I_0})$.

(2) The map from (1) is unique.

Indeed, by then comparing any two I and I' with a common one, we'd get the desired existence and *uniqueness* at the level of coefficients in A.

To prove (1), note that $\Lambda_I \in \widehat{\mathcal{C}}_\Lambda$ and θ_I deforms $\rho_{\widetilde{x}}$, and hence $\overline{\rho}$. Here is the picture:



The induced map $R \to \Lambda_I$ respects the map to Λ coming from the fact that $\rho_{\widetilde{x}}$ deforms $\overline{\rho}$, because if not, then we would have another map $R \to \Lambda_I \to \Lambda$, which contradicts the universal property of R.

To prove (2) just use the uniqueness from the universal property of (R, ρ^{univ}) for deforms on $\widehat{\mathcal{C}}_\Lambda$.

Lecture 5: Schlessinger's criterion and deformation conditions

Brandon Levin

October 30, 2009

1. What does it take to be representable?

We have been discussing for several weeks deformation problems, and we have said that we would like our deformation functors to be representable so we can study their ring-theoretic properties. We have stated that the framed deformation functor is always representable and that the unrestricted deformation functor is under certain hypotheses, but we have yet to prove either assertion.

There is a general theory of functors on C_Λ the category of Artin local Λ-algebras. My goal in this section is to, in as concrete terms as possible, describe what it takes for such a functor to be representable and how we might verify these properties. I then verify these properties for $D_{\bar{\rho}}$ with $\mathrm{End}_k(\bar{\rho}) = k$.

Along the way, I will point out some subleties of the relationship between C_Λ and $\hat{C}_\Lambda$.

Let k be a finite field. Recall that C_Λ is the category of local Artin Λ-algebras with residue field k, where Λ is any complete noetherian ring with residue field k. One can just think of $\Lambda = W(k)$ in which case every local Artin ring with residue field k admits a unique Λ-algebra structure. Denote by $\hat{C}_\Lambda$ the category of complete local Noetherian rings with residue field k.

We are interested in functors $F : C_\Lambda \to \mathrm{Set}$. We say F is representable if there exists $R \in \hat{C}_\Lambda$ such that F is naturally isomorphic to $\mathrm{Hom}_\Lambda(R, \cdot)$. (Technically you might call this pro-representable but it won't cause any confusion to just say "representable").

Elementary Properties of Representable Functors On C_Λ

I If F is representable, $F(k) = \mathrm{Hom}_\Lambda(R, k) =$ single point. We assume from now on that $F(k)$ is the one point set.

II If F is representable, $F(k[\epsilon]) = \mathrm{Hom}_k(m_R/m_R^2 + m_\Lambda, k) = t_F$ is a finite dimensional vector space over k.

III If F is representable, then F commutes with fiber products, i.e. if $A \to C$ and $B \to C$ are two maps in C_Λ then the natural map

$$F(A \times_C B) \to F(A) \times_{F(C)} F(B) \tag{1}$$

is a bijection.

Exercise 1.1. If you have not seen it before, you should verify that fiber products exist in C_Λ induced by set-theoretic fiber products. This is not true in $\hat{C}_\Lambda$ (See Conrad's example in Mazur's article **??**).

Exercise 1.2. Show that there is a natural multiplicative map $k \to End_\Lambda(k[\epsilon])$ given by $a \mapsto \alpha_a$ where $\alpha_a(x + y\epsilon) = x + ay\epsilon$.

Remark 1.3 (Tangent Space). Without knowing F is representable, its worth noting what is required for the tangent space to make sense. Let $t_F := F(k[\epsilon])$. The natural map $k \to End_\Lambda(k[\epsilon])$ induces scalar multiplication on t_F. We also know $k[\epsilon]$ is a group object in C_Λ compatible with scaling. Functoriality gives a map

$$F(k[\epsilon] \times_k k[\epsilon]) \to t_F.$$

If we can identify the LHS with $t_F \times t_F$, then we are set. The LHS does admit a natural map to $F(k[\epsilon]) \times_{F(k)} F(k[\epsilon]) = t_F \times t_F$. If this map is bijective (a special case of (III)), then t_F has a vector space structure.

Exercise: Check for F representable that the vector space structure on t_F given functorially as above is the same as the natural vector space structure on $\mathrm{Hom}_k(m_R/m_R^2 + m_\Lambda, k)$.

It turns out that property (III) along with the $\dim_k t_F < \infty$ is necessary and sufficient for F to be representable. However, I will not prove this because it is far too general to be useful. It could be quite difficult to check (III) for every possible pair of morphisms. Luckily, we don't have to! And this leads us to Schlessinger's criterion after a brief definition.

Definition 1.4. A map $A \to B$ in C_Λ is *small* if its surjective and its kernel is principal and annihilated by m_A.

Theorem 1.5 (Schlessinger's Criterion)**.** *Let F be a functor from C_Λ to Sets such that $F(k)$ is a single point. For any two morphism $A \to C$ and $B \to C$ consider the morphism* (1)*.*

Then if F has the following properties:

H1 (1) *is a surjection whenever $B \to C$ is small.*

H2 (1) *is a bijection when $C = k$ and $B = k[\epsilon]$.*

H3 *t_F is finite dimensional*

H4 (1) *is a bijection whenever $A \to C$ and $B \to C$ are equal and small.*

then F is representable.

I will return to the proof at the end given sufficient time, but one essentially makes clever use of the structure of Artin local rings working at each nilpotent level to build the representing ring as an inverse limit (see Schlessinger [**?**]. We denote the criterion by (SC).

Schlessinger's criterion is just one of many ways to show the deformation functor is representable. It has the advantage that it is concrete and allows one to really exploit the fact that we are working over Artin rings.

Proposition 1.6. *Assume that $End_G(\bar{\rho}) = k$ and G satisfies the p-finiteness condition. Then the deformation functor $D_{\bar{\rho}}$ is representable.*

Proof. I leave it to the reader to verify the following useful fact $\mathrm{Gl}_n(A \times_C B) \cong Gl_n(A) \times_{\mathrm{Gl}_n(C)} \mathrm{Gl}_n(B)$ as groups. This says that given any two lifts ρ_A and ρ_B which agree when pushed-forward to C come from a

lift to the fiber product. The only difficulty then in verifying H1 through H4 will be the ambiguity coming from conjugation.

In what follows ρ_A and ρ_B will always lifts of $\bar{\rho}$. We assume we have maps $A \to C$ and $B \to C$ satisfying the hypotheses in (SC). Further, we denote by $\tilde{\rho}_A$ and $\tilde{\rho}_B$ the respective push-forwards of ρ_A and ρ_B under the given maps.

(H1) We are given ρ_A and ρ_B such that $\tilde{\rho}_A = M\tilde{\rho}_B M^{-1}$ for some $M \in \mathrm{Gl}_n(C)$. Since $B \to C$ is surjective, we can lift M to $\mathrm{Gl}_n(B)$. Replacing ρ_B by $M\rho_B M^{-1}$ in the same deformation class yields compatible lifts which can then be lifted to $A \times_C B$.

(H2) Here we start with ρ and ρ' over the fiber product such that ρ_A and ρ_B are conjugate to ρ'_A and ρ'_B respectively. Choose conjugators M_A and M_B. Note that if $\tilde{M}_A$ and $\tilde{M}_B$ were equal, we could lift them to $\mathrm{Gl}_n(A \times_C B)$ and we would be done. This is true in general with no hypotheses on A, B, and C.

We are free to multiply M_B on the right by any matrix N such $N\rho_B N^{-1} = \rho_B$. Let $\mathrm{Stab}(\rho_B)$ be the set of such N. Further note that $\tilde{M}_B^{-1} * \tilde{M}_A$ is in $\mathrm{Stab}(\tilde{\rho}_B)$. If we can lift this to $\mathrm{Stab}(\rho_B)$ then we are done. Hence a sufficient condition for the desired map to be injective is that:

$$\mathrm{Stab}(\rho_B) \to \mathrm{Stab}(\tilde{\rho}_B) \tag{2}$$

is surjective.

This is clear since for $C = k$, $\mathrm{Stab}(\tilde{\rho}_B) = k^*$. I leave it as a exercise to show that for $B = k[\epsilon]$ and $C = k$, the equation (2) holds without any hypotheses on $\bar{\rho}$.

(H3) Follows from p-finiteness of Galois groups, given that (H2) implies the existence of the tangent space.

(H4) I leave it to the reader to verify that surjectivity of (2) follows from the following lemma:

If $\mathrm{End}_G(\bar{\rho}) = k$ and ρ_A is any lift of $\bar{\rho}$, then

$$\mathrm{End}_G(\rho_A) = A.$$

Set $L = \mathrm{End}_G(\rho_A)$ and note that L is an A-submodule of $\mathrm{End}(\rho_A)$ which contains the scalar matrices $A * I$. Further, we have that $L \times_A A/m_A = \mathrm{End}_G(\bar{\rho})$. By Nakayama, L is generated over k by any lift of I. Thus, $L = A$. □

If $\bar{\rho}$ is absolutely irreducible then it will satisfy the above hypothesis by Schur's lemma. However, there is important other case where $\bar{\rho}$ is not irreducible but still satisfies $\mathrm{End}_G(\bar{\rho}) = k$.

Proposition 1.7. *Let k be any field, and let V be any representation of G with a G-stable filtration $V_1 \subset V_2 \subset \ldots \subset V_n = V$ such that:*

I *V_{i+1}/V_i is one-dimensional with G acting by χ_i.*

II *The χ_i are distinct.*

III *The extension $V_i/V_{i-1} \to V_{i+1}/V_{i-1} \to V_{i+1}/V_i$ is non-split for all i.*

Then $\mathrm{End}_G(V) = k$.

Proof. Before you read this proof, I recommend doing the 2 by 2 case by hand which I may or may not have gotten to in the lecture.

Let $M \in \mathrm{End}_G(V)$ i.e. M commutes with the G-action. We want to show that M is a scalar. We first note that V_1 is the unique 1-dimensional subspace on which G acts via χ_1. For if V_1' were another, we could build a Jordan-Holder series $V_1 \subset V_1 \cup V_1' \subset \dots$ and thus χ_1 would appear at least twice in Jordan-Holder decomposition which can't happen since χ_i are distinct.

It follows then that M preserves V_1 and by induction the whole flag. Let M act on V_1 by multiplication by a. We claim that $M = aI$. Consider $M - aI : V \to V$ also in $\mathrm{End}_G(V)$. Since $M - aI|_{V_1} = 0$, it factors as a morphism

$$T : V/V_1 \to V.$$

By induction, the induced map $V/V_1 \to V/V_1$ which is G-invariant is multiplication by a scalar c. If $c \neq 0$, then $T|_{V_2}$ would give a splitting of the extension where $i = 1$ and so we can assume $c = 0$.

Thus, T is actually a G-invariant map

$$V/V_1 \to V_1.$$

If $T = 0$ we are done, else let V_i be the first subspace on which it is non-trivial. Then $T : V_i/V_{i-1} \to V_1$ is an isomorphism as G-modules, contradiction. □

Remark 1.8. Schlessinger's criterion is a statement purely about a functor on C_Λ. However, once we know F is representable, its quite natural to talk about its points valued in complete local Noetherian rings for we have $\mathrm{Hom}_\Lambda(R, A) = \mathrm{Hom}_\Lambda(R, \lim A/{m_A}^n) = \lim F(A/{m_A}^n)$. In fact, that's really what we were interested in all along, for example, $\mathbb{Z}_p$-deformations, not representations on Artin local rings. So we must ask ourselves, are the points of our universal deformation ring valued in $\hat{C}_\Lambda$ what we want them to be?

Let A be a complete local Noetherian ring. Its clear the any deformation to A yields a map $R_{\bar\rho}^{\mathrm{univ}} \to A$ (here's where you use that your representation is continuous). However, the other direction requires an argument. Denote A/m_A^n by A_n. We are given deformations $\rho_n \in D_{\bar\rho}(A_n)$ such that $\rho_n \otimes A_{n-1}$ is equivalent to ρ_{n-1}. If ρ_n formed a compatible system of lifts, we would be fine, but we have conjugations interfering at each level. In this case, it can be resolved quite easily. Assume we have compatibility up to ρ_n. Given that $M(\rho_{n+1} \otimes A_n)M^{-1} = \rho_n$ change ρ_{n+1} by any lift of M to $\mathrm{Gl}_N(A_n)$ and proceed by induction.

We will return to this point again later where the argument will require some extra input.

Remark 1.9. (Framed Deformations) The fact mentioned earlier that $\mathrm{Gl}_n(A \times_C B) \cong Gl_n(A) \times_{\mathrm{Gl}_n(C)} \mathrm{Gl}_n(B)$ implies that the framed deformation functor $D_{\bar\rho}^{\square}$ commutes with all fiber products and thus is representable. However, (SC) is probably way to fancy a way to prove existence for framed deformations. For the record, I give a proof that $R_{\bar\rho}^{\square}$ exists.

Proof. Since $\mathrm{Gl}_N(k)$ is finite, $\bar\rho$ is trivial on some finite index subgroup H of G. For any lift, we have that

$$\rho_A(H) \subset \ker(\mathrm{Gl}_N(A) \to \mathrm{Gl}(k))$$

which is a p-group for any $A \in C_\Lambda$. Thus, $\rho_A|_H$ factors through maximal pro-p quotient of H which by p-finiteness is topologically finitely generated. Pick generators $g_1, \ldots, g_j$. Also, pick coset representatives for $g_{j+1}, \ldots, g_m$ for G/H.

Any lift ρ_A is determined by where the $\{g_i\}$ are sent. Consider the power series ring $R = \Lambda[[X_{i,j}^l]]$ where $1 \leq i, j \leq N$ and $1 \leq l \leq m$. I claim that we can construct the universal framed deformation ring as a quotient of R such that the universal framed deformation ρ^{univ} is given by g_l goes to the matrix $(X_{i,j}^l)$. Let S be the set of relations in G amongst the g_l. For any relation, we can consider the corresponding relation on matrices under the map g_l goes to $(X_{i,j}^l)$. We form the ideal I in R generated by these relations. Then $D_{\bar{\rho}}^{\square}$ is represented by R/I. □

Before we move on to deformation conditions, I would like to recall several different interpretation of the tangent space which will be useful in the future. We would like to give a concrete interpretation of the abstract $t_{\bar{\rho}} := D_{\bar{\rho}}(k[\epsilon])$. Let (V, τ) be a deformation to $k[\epsilon]$ where τ is an isomorphism $V/\epsilon \to \bar{\rho}$. Consider the following exact sequence of k-vector spaces:

$$0 \to \epsilon V \to V \to V/\epsilon V \to 0.$$

Simply because G commutes with the action of $k[\epsilon]$, this is an extension of G-modules. Further one can identify via τ the terms on both ends with $\bar{\rho}$.

Hence we get a map $t_{\bar{\rho}} \to \text{Ext}(\bar{\rho}, \bar{\rho})$. It is an exercise to show the map is bijective. By general nonsense, one can identify this Ext-group with $H^1(G, \text{ad}(\bar{\rho}))$. Note that $\text{ad}(\bar{\rho})$ is just $\text{End}(\bar{\rho})$ where G acts via conjugation (ad stands for adjoint).

I will give you the map $\text{Ext}(\bar{\rho}, \bar{\rho})$ to $H^1(G, \text{ad}(\bar{\rho}))$, but I leave it to you to check that the vector space structures on $t_{\bar{\rho}}$ and $H^1(G, \text{ad}(\bar{\rho}))$ agree.

Given an extension $0 \to V' \to V \to V'' \to 0$ choose a splitting $\phi : V'' \to V$ just as vector spaces. The map $g \mapsto g\phi g^{-1} - \phi$ is a co-cycle with values in $\text{ad}(\bar{\rho})$.

In the next section, as we impose various deformation conditions, we will eventually want to keep track of the effect on the tangent space.

2. Deformation Conditions

As we have mentioned already several times in this seminar, whether in the local or global situation, the unrestricted universal deformation ring if it exists will be far too "big" to be useful. Hence we will want to impose some conditions on what kinds of deformations we allow. Deformation conditions can come in different varieties. Often we have global representations on which we impose local conditions at finite set of primes. At these local places, we might impose matrix conditions for example, fixed determinant, ordinary, etc. We could also impose conditions coming from geometry or p-adic Hodge theory: flat, crystalline, semi-stable. I will discuss some of these in more detail later.

In Mazur's article [**?**], he defines the notion of deformation condition quite generally such that everything we will talk about probably fits into that framework. However, for our purposes and for the purpose of

intuition, the definitions are unilluminating. Instead, I will give two different perspective from which one could derive all the definitions.

Functorial Perspective If $D_{\bar{\rho}}$ is the deformation functor, a deformation condition should define a subfunctor $D'_{\bar{\rho}}$ of $D_{\bar{\rho}}$. Further, if $D_{\bar{\rho}}$ is representable, then $D'_{\bar{\rho}}$ should be as well. One could use the term relatively representable as Mok did, but its not necessary.

The first statement is usually immediate for any deformation condition. The second one is not. This is a place you might use Schlessinger's criterion, maybe you know already that $D_{\bar{\rho}}$ satisfies Schlessinger then you just have to show the $D'_{\bar{\rho}}$ does too. We will see an example of this soon.

Deformation Space Perspective If $D_{\bar{\rho}}$ is represented by $R_{\bar{\rho}}$, then we can talk about Spec $R_{\bar{\rho}}$ as the space of all deformations of $\bar{\rho}$. Personally, I find this picture quite compelling.

A brief aside. Say $\bar{\rho}$ is modular, then Akshay explained that we get a surjective map $R_{\bar{\rho}} \to T$, a Hecke ring. In geometric language, Spec T is a closed subspace of the deformation space which includes the closed point corresponding to $\bar{\rho}$. Imagine this as the locus of "modular" deformations. Given a representation coming from an elliptic curve, etale cohomology, or somewhere else, whose reduction is $\bar{\rho}$, its natural to ask does it land in that locus. Our goal then, as I understand it, is to impose enough purely representation theoretic condition to cut out the "modular" locus. Then whatever representation we started with will presumably have those properties and hence will be modular.

From this perspective then a deformation condition is just a closed condition on the space of all deformations. More concretely, there exist an ideal I such that for any $f : R \to A$, $f \circ \rho_{\text{univ}}$ satisfies the deformation condition if and only if f factors through R/I.

Remark 2.1. We can connect the two perspectives as follows: let D' be subfunctor of D and assume they are both representable by R and R', then we get a natural map $R \to R'$. I claim this map is surjective. It suffices to check that the map on cotangent spaces is surjective. But the map on cotangent spaces is dual to the map on tangent spaces which is injective because $D'(k[\epsilon]) \subset D(k[\epsilon])$.

Remark 2.2 (Relative Perspective). There is relative perspective which doesn't require $D_{\bar{\rho}}$ to be representable. Given any deformation ρ to A, we can ask if the subset of maps Spec $B \to$ Spec A such that the pullback of ρ has a given condition is represented by a closed subset of Spec A? If this holds for all A and if the univeral deformation ring exists then we can apply it to $(R_{\bar{\rho}}, \rho^{\text{univ}})$. This is the perspective Kisin often takes.

2.1. **Determinant Condition.** Let G be any local or global Galois group.

Definition 2.3. Let $A \in \hat{C}_\Lambda$, and let $\delta : G \to \Lambda^*$ be a character. We say a representation ρ on a free rank n A-module has *determinant* δ if $\wedge^n : G \to A^*$ factors through δ.

Consider the functor of deformations with fixed determinant δ (assume that $\bar{\rho}$ has determinant δ). I claim this is a deformation condition.

Here the second perspective is most natural. Let $\delta^{\text{univ}} : G \to R^*_{\bar\rho}$ be the the determinant of the universal deformation of $\bar\rho$ assuming it exists. Then let I be the ideal generated by $\delta^{\text{univ}}(g) - i(\delta(g))$ where $i : \Lambda^* \to R^*_{\bar\rho}$ is inclusion coming from algebra structure. Then, $R_{\bar\rho}/I$ represents deformations with determinant δ. Its usually denoted by $R^\delta_{\bar\rho}$.

I haven't worked it out, but I suspect it would quite a bit more tedious to show for example that the determinant condition defines a subfunctor which satisfies Schlessinger's criterion or that it is relatively representable.

Note that even if $D_{\bar\rho}$ is not representable, the same proof goes through for any (A, ρ) to show relative closedness as in Remark 2.2.

2.2. **Unramified Condition.** Let K be a global field and let S be a finite set of primes. We denote by $G_{K,S}$ the maximal Galois group unramified outside S. Take $\bar\rho$ to be a residual representation of $G_{K,S}$ which happens also to be unramified at some $\nu \in S$.

Definition 2.4. Let ρ_A be any deformation of $\bar\rho$. We say that ρ_A is unramified at ν if $\rho_A|_{G_{K_\nu}}$ is unramified for any choice of decomposition group G_{K_ν}.

In showing this is a deformation condition, I will illustrate the relative perpective. Again, let ρ_A be any deformation of $\bar\rho$. Consider any map $f : A \to B$. The push-forward $f_*(\rho_A)$ will be unramified an ν iff its trivial on the inertia group I_{K_ν} (for some choice of inertia).

Let J be the ideal in A generated by the entries of $\{I - \rho_A(g)\}$ for all $g \in I_{K_\nu}$. Then, one can verify that A/J represents the unramified at ν condition. This is the relative condition; if $R_{\bar\rho}$ exists, we can apply the same argument to construct the universal deformation ring unramified at ν.

2.3. **Ordinary Deformations.** We will go into extensive detail in this section as the notion of ordinary will play a prominent role in what is to come. There seem to be several definitions of ordinary floating around. I chose one that is both concrete and sufficiently general for now.

Definition 2.5. Let $G = G_K$ be a local Galois group where the residue characteristic is p. Let $\psi : G \to \mathbb{Z}_p^*$ be the p-adic cyclotomic character. An n-dimensional representation ρ of G is *ordinary* if

$$\rho|_{I_K} \sim \begin{pmatrix} \psi^{e_1} & \star & \star & \star \\ 0 & \psi^{e_2} & \star & \star \\ 0 & 0 & \ddots & \star \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where $e_1 > e_2 > \ldots > e_{n-1} > 0$. Implicitly we are including $\mathbb{Z}_p \to \Lambda$ so that the definition makes sense over any Λ-algebra and hence on our category $\hat{C}_\Lambda$.

Before we continue, let me say where this condition is coming from.

Example 2.6. If E is an elliptic curve over K a local field of residue characteristic p which has good ordinary reduction at p, then the representation of G_K on $T_p(E)$ is ordinary. In particular, it has the form

$$g \mapsto \begin{pmatrix} \psi(g)\chi_1(g) & \star \\ 0 & \chi_2(g) \end{pmatrix}$$

where χ_1 and χ_2 are unramified characters on G_K.

There is an corresponding notion of what it means for a modular form to be ordinary, but I won't get into it here.

Remark 2.7. Note that though ψ^e is non-trivial on I_K for any $e \neq 0$ (in fact its infinitely ramified), its possible for ψ^e to be residually trivial. The residually trivial case will be of interest to us later on, but for now, we assume that all ψ^{e_i} are residually distinct and non-trivial. Its not hard to work out exactly when this happens based on K, p, e_i.

Next, we would like to show that if we conjugate ρ such that $\rho|_{I_K}$ is upper triangular, then ρ will be upper triangular. This follows from the following useful lemma.

Lemma 2.8. *Let $\rho : I_K \to \mathrm{Gl}_N(A)$ be a representation, $A \in C_\Lambda$ landing in the upper triangular matrices with residually distinct characters along the diagonal. If $M\rho M^{-1}$ is also upper triangular with same characters occurring in the same order, then M is upper triangular.*

Proof. We prove it in two steps. First we show that ρ preserves a unique flag. We know this fact residually using a Jordan-Holder component argument as in Proposition 1.7. Let $M \cong A^N$ with G acting through ρ. Let $L_1 \subset M$ be the line corresponding to e_1 on which ρ acts by χ_1. We want to show that given any $m \in M$ on which G acts via χ_1, $m \in L_1$. From there, it is a simple induction on N.

Consider the quotient $N = M/L_1$. I claim N contains no non-zero v on which G acts via χ_1. Assume there existed such an v. Filter N by $m_A^n N$, and let n_0 be the smallest n such that $v \notin m_A^n N$. Clearly G acts on the image of v in $m_A^{n_0-1}N/m_A^{n_0}N$ via the character χ_1. However, its not hard to see that for any n,

$$m_A^n N/m_A^{n+1}N \cong m_A^n/m_A^{n+1} \otimes_k (\bar{\rho}/(L_1 \otimes k))$$

as $k[G]$-modules. The RHS breaks up as the direct sum of copies of $\bar{\rho}$ quotiented by the χ_1 subspace and hence χ_1 doesn't appear anywhere in semi-simplification (using residual distinctness). Hence the flag is unique.

To say M is upper triangular is equivalent to saying M preserves the flag which we have now shown to be unique. Again, by induction on N, it will suffice to show the M preserves L_1. Let $e_1 \in L_1$. Our hypotheses imply that

$$M\rho M^{-1}e_1 = \rho e_1 = \chi_1 e_1.$$

Multiplying by M^{-1} and using that M^{-1} commutes with χ_1, we get

$$\rho(M^{-1}e_1) = \chi(M^{-1}e_1).$$

By uniqueness then, $M^{-1}e_1 \in L_1$ and hence M^{-1} preserves the flag so M does as well. □

Corollary 1. *If ρ is ordinary, then it lands in a Borel subgroup, i.e. is upper-triangular with respect to some basis.*

Proof. By assumption, we can conjugate $\rho|_I$ to be upper-triangular so it suffices to show that $\rho(g)$ is, where g is some Frobenius element. Since ψ^e are invariant under conjugation by g, we see that $\rho(g)$ satisfies the hypotheses on M in the previous lemma and hence is upper-triangular. □

Corollary 2. *The ordinary deformation functor satisfies (SC) and so is representable under the assumption that the residual representation is non-split in the sense of Proposition 1.7.*

Proof. By Proposition 1.7, the residual deformation satisifies the necessary conditions for the universal deformation functor to exist. We denote the ordinary deformation functor by $D_{\bar{\rho}}^{\mathrm{ord}}$; it is clearly a subfunctor of $D_{\bar{\rho}}$. As a subfunctor, injectivity of the map (1) is automatic in H1, H2, and H4. Hence it suffices to check that (1) is surjective under the hypotheses of H1, namely when $B \to C$ is small.

We are free to choose ordinary lifts ρ_A and ρ_B

$$\tilde{\rho}_A = M\tilde{\rho}_B M^{-1}.$$

Since $\tilde{\rho}_A$ and $\tilde{\rho}_B$ are both ordinary M satisfies the hypotheses of the previous lemma and so is upper-triangular. We can choose a lift M' of M to $\mathrm{Gl}_n(B)$ which is upper triangular. Changing ρ_B by M' maintains its ordinary form. Hence, we have ρ_A and ρ_B agreeing after push-forward and both have ordinary form and so their fiber product will also be ordinary. □

Exercise 2.9. (Continuity) Let A be complete local Noetherian ring and set $A_n = A/m_A^n$. Given a compatible system of ordinary deformations ρ_n, show that there exist an ordinary deformation ρ_A such that $\rho_A \otimes A_n \equiv \rho_n$. Hint: See Remark 1.8.

As a final comment, it is possible to interpret ordinarity as a closed condition at least under the assumptions of residually distinct and nonsplit, but I did not have to write it up. Hopefully, I will have a chance to present it in seminar. Otherwise, feel free to ask me about it afterwards.

Lecture 6: Presentations of deformation rings

Samit
November 6, 2009
Notes by Sam Lichtenstein

This lecture is about getting bounds for the dimension of deformation rings, by bounding the number of generators and relations. The reference for this lecture is Kisin's article in CDM, or stuff from his Hawaii notes.

1. Local setup and statement

Let $K/\mathbb{Q}_p$ be finite, $\mathcal{O} = \mathcal{O}_K, \pi$ a uniformizer, $k = \mathcal{O}/(\pi)$, Γ a profinite group satisfying the p-finiteness condition "Φ_p", and $\overline{\rho} : \Gamma \to \mathrm{GL}_n(k)$ a mod π representation. We consider deformations to complete local noetherian $\mathcal{O}$-algebras with residue field k. The framed deformation ring $R^{\square}_{\overline{\rho}}$ always exists, so we have a universal representation

$$\Gamma \overset{\rho^{\square}_{\mathrm{univ}}}{\to} \mathrm{GL}_n(R^{\square}_{\overline{\rho}}).$$

Assuming $\mathrm{End}_{\Gamma}\,\overline{\rho} = k$, we also know $R_{\overline{\rho}}$ exists, and we then get a universal deformation

$$\Gamma \overset{\rho^{\mathrm{univ}}}{\to} \mathrm{GL}_n(R_{\overline{\rho}}).$$

Recall that

$$D^{\square}_{\overline{\rho}}(k[\epsilon]) = \mathrm{Hom}_k(\mathfrak{m}_{R^{\square}}/(\mathfrak{m}^2_{R^{\square}}, \pi), k) \cong Z^1(\Gamma, \mathrm{ad}\,\overline{\rho})$$

and $D_{\overline{\rho}}(k[\epsilon]) = \mathrm{H}^1(\Gamma, \mathrm{ad}\,\overline{\rho})$ as k-vector spaces.

Theorem. *Let $r = \dim_k Z^1(\Gamma, \mathrm{ad}\,\overline{\rho})$. Then there exists an $\mathcal{O}$-algebra isomorphism*

$$\mathcal{O}[\![x_1, \ldots, x_r]\!]/(f_1, \ldots, f_s) \cong R^{\square}_{\overline{\rho}}$$

where $s = \dim_k \mathrm{H}^2(\Gamma, \mathrm{ad}\,\overline{\rho})$.

Corollary. (i) $\dim R^{\square}_{\overline{\rho}} \geq 1 + n^2 - \chi(\Gamma, \mathrm{ad}\,\overline{\rho}) = 1 + n^2 - h^0(\mathrm{ad}\,\overline{\rho}) + h^1(\mathrm{ad}\,\overline{\rho}) - h^2(\mathrm{ad}\,\overline{\rho})$.
(ii) $\dim R_{\overline{\rho}} \geq 2 - \chi(\Gamma, \mathrm{ad}\,\overline{\rho})$.

Proof of corollary. From $\mathcal{O}$ we get a contribution of 1. hence we get $\dim R^{\square}_{\overline{\rho}} \geq 1 + \dim Z^1 - h^2$. Now (i) follows formally noting that $\dim Z^0 = \dim C^0 = n^2$. (Use $h^1 = \dim Z^1 - \dim B^1$ and $\dim B^1 = \dim C^0 - \dim Z^0 = \dim C^0 - h^0$.) Then (ii) is immediate using the fact that $R^{\square}_{\overline{\rho}}$ is basically a PGL_n-bundle over $R_{\overline{\rho}}$. □

2. Proof of Theorem 1

Using completeness [exercise] we can choose a surjection

$$\varphi : \mathcal{O}[\![x]\!] := \mathcal{O}[\![x_1, \ldots, x_r]\!] \twoheadrightarrow R^{\square}_{\overline{\rho}}.$$

(Send the x_i's to elements which reduce to a basis for the tangent space $Z^1(\Gamma, \mathrm{ad}\,\overline{\rho})$ of the framed deformation ring.) The problem is to show that the minimal number of generators of the kernel $\mathbb{J} = \ker\varphi \subset \mathcal{O}[\![x]\!]$ is at most s. Let $\mathbf{m} = \mathfrak{m}_{\mathcal{O}}[\![x]\!] \subset \mathcal{O}[\![x]\!]$ be the maximal ideal $(\pi, x_1, \ldots, x_r)$. It would suffice to construct a linear injection $(\mathbb{J}/\mathbf{m}\mathbb{J})^* \hookrightarrow \mathrm{H}^2(\Gamma, \mathrm{ad}\,\overline{\rho})$. There is a subtle technical problem in an attempt to construct such an injection. We explain the problem, and then the fix to get around it.

For each $\gamma \in \Gamma$ choose a set-theoretic lift $\widetilde{\rho}(\gamma) \in \mathrm{GL}_n(\mathcal{O}[\![x]\!]/\mathbf{m}\mathbb{J})$ of $\rho^{\square}(\gamma) \in \mathrm{GL}_n(\mathcal{O}[\![x]\!]/\mathbb{J}) = \mathrm{GL}_n(R^{\square})$. We need to make this choice so that $\widetilde{\rho}$ is a *continuous* function of γ. It is not clear if the map

$$\mathcal{O}[\![x]\!]/\mathbf{m}\mathbb{J} \twoheadrightarrow \mathcal{O}[\![x]\!]/\mathbb{J}$$

admits a continuous section as topological spaces, so it is not clear how to find a continuous $\widetilde{\rho}$. To handle this problem, we now prove:

Claim: For $r > 0$, let $\mathbb{J}_r = (\mathbb{J} + \mathbf{m}^r)/\mathbf{m}^r \in \mathcal{O}[\![x]\!]/\mathbf{m}^r$ and let $\mathbf{m}_r = \mathbf{m}/\mathbf{m}^r$. For $r \gg 0$, the natural map $\mathbb{J}/\mathbf{m}\mathbb{J} \to \mathbb{J}_r/\mathbf{m}_r\mathbb{J}_r$ is an isomorphism.

Proof. The map is surjective, and for injectivity we have to show that $\mathbb{J} \cap (\mathbf{m}\mathbb{J} + \mathbf{m}^r) = \mathbf{m}\mathbb{J}$ for large r. Certainly $\mathbf{m}\mathbb{J}$ lies in the intersection for all r, so since $\mathbb{J}/\mathbf{m}\mathbb{J}$ has finite length we see that the intersection stabilizes at some intermediate ideal for $r \gg 0$. This stabilizing ideal must then be the total intersection. But by Artin-Rees applied to $\mathbf{m}\mathbb{J}$ as a finite $\mathcal{O}[\![x]\!]$-module, the intersection of all $(\mathbf{m}\mathbb{J} + \mathbf{m}^r)$'s is $\mathbf{m}\mathbb{J}$. $\square$

By the Claim, to prove the desired result about minimal number of generators of $\mathbb{J}$, we can replace $\mathcal{O}[\![x]\!]$ and $R^\square := R^\square_{\overline{\rho}}$ with their quotients by rth power of maximal ideal for some large r. The quotient of $R^\square$ by rth power of its maximal ideal is universal in the category of complete local noetherian $\mathcal{O}$-algebras whose maximal ideal has vanishing rth power (exercise!). So working within this full subcategory of local $\mathcal{O}$-algebras, we can still exploit universal mapping properties. But we gain the advantage that now our rings are of finite length as $\mathcal{O}/\pi^r$-modules, so in particular they're all discrete with their max-adic topology and hence the Galois representations which arise have open kernel. We can therefore find the required continuous section, working throughout with local rings whose maximal ideal has a fixed but large order of nilpotence.

So we now proceed in such a modified setting (so the definition of $\mathbb{J}$ changes accordingly, but the Claim shows that this does not affect $\mathbb{J}/\mathbf{m}\mathbb{J}$, which is to say the minimal number of generators of $\mathbb{J}$). In particular, in the new setting we will construct a k-linear injection of $\mathbb{J}/\mathfrak{m}\mathbb{J}$ into $\mathrm{H}^2(\Gamma, \mathrm{ad}\,\overline{\rho})$, thereby finishing the proof.

For $f \in (\mathbb{J}/\mathbf{m}\mathbb{J})^*$. let

$$\rho_f(\gamma, \delta) = f(\widetilde{\rho}(\gamma\delta)\widetilde{\rho}(\delta)^{-1}\widetilde{\rho}(\gamma)^{-1} - \mathbf{1}),$$

where we apply the map f "entry-wise" to the given matrix in $\mathrm{Mat}_{n\times n}(\mathbb{J}/\mathbf{m}\mathbb{J})$. That is, the map φ_f has the form

$$\Gamma^2 \to \mathrm{Mat}_{n\times n}(\mathbb{J}/\mathbf{m}\mathbb{J}) \xrightarrow{f} \mathrm{Mat}_{n\times n}(k).$$

Now we observe the following facts.

(1) $\varphi_f \in Z^2(\Gamma, \mathrm{ad}\,\overline{\rho})$.
(2) $[\varphi_f] \in \mathrm{H}^2(\Gamma, \mathrm{ad}\,\overline{\rho})$ is independent of the choice of lift $\widetilde{\rho}$.
(3) $f \mapsto [\varphi_f]$ is k-linear.
(4) $f \mapsto [\varphi_f]$ is injective, but more precisely we have $[\varphi_f] = 0 \Leftrightarrow$ we can choose $\widetilde{\rho}$ to be a homomorphism "mod $\mathbb{J}_f$" where $\mathbb{J}_f = \ker(\mathbb{J} \to \mathbb{J}/\mathbf{m}\mathbb{J} \xrightarrow{f} k) \Leftrightarrow f = 0 \Leftrightarrow \mathbb{J}_f = \mathbb{J}$.

Note that (4) provides the desired linear injection, and hence proves the theorem; (1)-(3) are necessary to make sense of (4).

Let us prove the facts above.

(1) This is a formal computation, which goes as follows. Note that we can identify $\mathrm{Mat}_{n\times n}(\mathbb{J}/\mathbf{m}\mathbb{J})$ under addition with $(\mathbf{1} + \mathrm{Mat}_{n\times n}(\mathbb{J}/\mathbf{m}\mathbb{J}))$ under multiplication, since $\mathbb{J} \subset \mathbf{m}$. Using this identification, we have

$$\mathrm{d}\,\varphi_f(\gamma, \delta, \epsilon) = \gamma\varphi_f(\delta, \epsilon) - \varphi_f(\gamma\delta, \epsilon) + \varphi_f(\gamma, \delta\epsilon) - \varphi_f(\gamma, \delta) \in \mathrm{Mat}_{n\times n}(k).$$

If we want to prove this is zero, it's enough to check "upstairs" in $\mathrm{Mat}_{n\times n}(\mathbb{J}/\mathbf{m}\mathbb{J})$, i.e. before applying f. Thus we really want to check that

$$(\widetilde{\rho}(\gamma)\widetilde{\rho}(\delta\epsilon)\widetilde{\rho}(\epsilon)^{-1}\widetilde{\rho}(\delta)^{-1}\widetilde{\rho}(\gamma)^{-1}) \times (\widetilde{\rho}(\gamma\delta)\widetilde{\rho}(\epsilon)\widetilde{\rho}(\gamma\delta\epsilon)^{-1}) \times (\widetilde{\rho}(\gamma\delta\epsilon)\widetilde{\rho}(\delta\epsilon)^{-1}\widetilde{\rho}(\gamma)^{-1}) \times (\widetilde{\rho}(\gamma)\widetilde{\rho}(\delta)\widetilde{\rho}(\gamma\delta)^{-1}) \stackrel{?}{=} \mathbf{1}\,.$$

The trick is to insert the bracketed term (which is 1) below:

$$\widetilde{\rho}(\gamma)\widetilde{\rho}(\delta\epsilon)\widetilde{\rho}(\epsilon)^{-1}\widetilde{\rho}(\delta)^{-1}\widetilde{\rho}(\gamma)^{-1}\widetilde{\rho}(\gamma\delta)\overbrace{\widetilde{\rho}(\delta)^{-1}\widetilde{\rho}(\delta)}^{insert}\widetilde{\rho}(\epsilon)\widetilde{\rho}(\gamma\delta\epsilon)^{-1} \times \widetilde{\rho}(\gamma\delta\epsilon)\widetilde{\rho}(\delta\epsilon)^{-1}\widetilde{\rho}(\gamma)^{-1} \times (\widetilde{\rho}(\gamma)\widetilde{\rho}(\delta)\widetilde{\rho}(\gamma\delta)^{-1}) \stackrel{?}{=} \mathbf{1}\,.$$

Now observe that the bracketed terms below reduce to 0 in $\mathrm{Mat}_{n\times n}(k)$ and hence can be commuted with one another (!):

$$\widetilde{\rho}(\gamma)\overbrace{\widetilde{\rho}(\delta\epsilon)\widetilde{\rho}(\epsilon)^{-1}\widetilde{\rho}(\delta)^{-1}}^{I}\overbrace{\widetilde{\rho}(\gamma)^{-1}\widetilde{\rho}(\gamma\delta)\widetilde{\rho}(\delta)^{-1}}^{II}\widetilde{\rho}(\delta)\widetilde{\rho}(\epsilon)\widetilde{\rho}(\gamma\delta\epsilon)^{-1} \times \widetilde{\rho}(\gamma\delta\epsilon)\widetilde{\rho}(\delta\epsilon)^{-1}\widetilde{\rho}(\gamma)^{-1} \times (\widetilde{\rho}(\gamma)\widetilde{\rho}(\delta)\widetilde{\rho}(\gamma\delta)^{-1}) \stackrel{?}{=} \mathbf{1}\,.$$

After swapping I and II one sees that in fact everything cancels magically. (Is there is a "conceptual" proof of (1)?)

(2) This is similar to (1). First write $\widetilde{\rho}^{\text{new}}(\gamma) = a(\gamma)\widetilde{\rho}(\gamma)$ for some

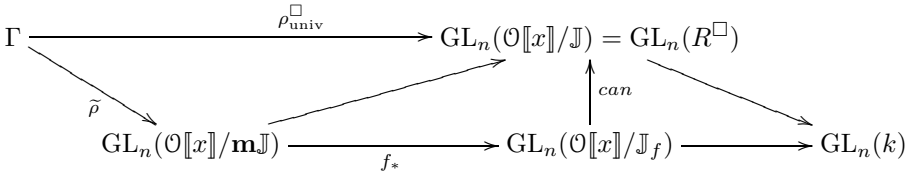
$$a : \Gamma \to \mathbf{1} + \mathrm{Mat}_{n\times n}(\mathbb{J}/\mathbf{m}\mathbb{J}).$$

The idea is to show formally that $a(\gamma)$ (which is of course a *continuous* 1-cocycle on Γ) changes φ_f by $\mathrm{d}\, a$. This is done with a similar "insert $\mathbf{1}$ cleverly and commute stuff" trick as in (1).

(3) OK.

(4) The last equivalence in (4) is clear. For the other two equivalences, the implications "$\Leftarrow$" are OK. The implication that $[\varphi_f] = 0$ implies we can choose $\widetilde{\rho}$ to be a homomorphism mod $\mathbb{J}_f$ follows from the previous calculation [omitted] that $\widetilde{\rho} \rightsquigarrow a \cdot \widetilde{\rho}$ changes φ by $\mathrm{d}\, a$. In particular, if φ is already a coboundary, then by changing the choice of lift we can make $\varphi = 0$, which is the same as saying our lift is a homomorphism mod $\mathbb{J}_f$. So the crux of the matter is the second "$\Rightarrow$".

Here's the situation. We have a diagram



We'd like to prove that $\mathcal{O}[\![x]\!]/\mathbb{J}_f \to \mathcal{O}[\![x]\!]/\mathbb{J}$ is an isomorphism. By the universality of $R^{\square}$ we get the map

$$\mathcal{O}[\![x]\!]/\mathbb{J} \overset{\exists!}{\to} \mathcal{O}[\![x]\!]/\mathbb{J}_f \overset{\text{can}}{\to} \mathcal{O}[\![x]\!]/\mathbb{J}$$

and again by universality the composition is the identity. Now it would be enough to check that $\mathbb{J} \subset \mathbb{J}_f$. Note that the image of x_i in $\mathcal{O}[\![x]\!]/\mathbb{J}$ maps to $x_i + a_i \in \mathcal{O}[\![x]\!]/\mathbb{J}_f$ where a_i is some element of $\mathbb{J}$. It will suffice to show that if $g(x_1, \ldots, x_n) \in \mathbb{J}$ then g maps to g itself in $\mathcal{O}[\![x]\!]/\mathbb{J}_f$.

First we claim that $\mathbb{J} \subset (\mathbf{m}^2, \pi)$ [recall that $\mathbb{J} = \ker(\mathcal{O}[\![x]\!] \twoheadrightarrow R^{\square})$]. Indeed, if $g \in \mathbb{J}$ then $g = g_0 + \sum g_i x_i + O(\mathbf{m}^2)$. Moreover $g_0 \in (\pi)$ and each g_i lies in (π) since the x_i's map to a *basis* of $\mathbf{m}/(\mathbf{m}^2, \pi)$. Thus $g \in (\mathbf{m}^2, \pi)$. Consequently, it's enough to show what we want for $g \in (\mathbf{m}^2, \pi)$. [This will be important later on!]

But if $g \in (\mathbf{m}^2, \pi)$ then under $\mathcal{O}[\![x]\!]/\mathbb{J} \to \mathcal{O}[\![x]\!]/\mathbb{J}_f$ we still have

$$g = g_0 + \sum g_i x_i + O(\mathbf{m}^2) \mapsto g_0 + \sum g_i(x_i + a_i) + O(\mathbf{m}^2),$$

and the observation is that when we subtract off g from this we get $\sum g_i a_i$ in the $O(\mathbf{m})$ term, which [by inspection] is in $\mathbf{m}\mathbb{J} \subset \mathbb{J}_f$. Similarly one sees that the higher order terms vanish mod $\mathbb{J}_f$.

This concludes the proof of (4), hence the claim, hence the theorem.

3. Completed tensor products

Example. Let R be a Noetherian ring, and consider $R[x] \otimes_R R[y] \cong R[x, y]$. However $R[\![x]\!] \otimes_R R[\![y]\!]$ is something weird, being just a part of $R[\![x, y]\!]$. It's easy to see that it does at least inject into $R[\![x, y]\!]$. The idea is that $M \otimes R^I \hookrightarrow M^I$ for any free R-module R^I (here I is an arbitrary index set) but this map fails to be an isomorphism.

To check the injectivity, note that it's OK for M finite free, which allows one to deduce it for M finitely presented, and then pass to a direct limit to conclude the general case. Applying this to $I = \mathbb{Z}$ and $M = R[\![x]\!]$ gives what we want in our case. But to see that our map $R[\![x]\!] \otimes R[\![y]\!] \hookrightarrow R[\![x, y]\!]$ is not surjective, observe that $\sum x^n y^n$ is not in the image!

Definition. Let $\mathcal{O}$ be a complete Noetherian local ring and R, S complete Noetherian local $\mathcal{O}$-algebras (meaning the structure maps are local morphisms). Assume at least one of the residue field extensions $\mathcal{O}/\mathfrak{m}_{\mathcal{O}} \subset R/\mathfrak{m}_R$ and $\mathcal{O}/\mathfrak{m}_{\mathcal{O}} \subset S/\mathfrak{m}_S$ is finite. Then set $\mathfrak{m} \lhd R \otimes_{\mathcal{O}} S$ to be the ideal generated by

$$\mathfrak{m}_R \otimes_{\mathcal{O}} S + R \otimes_{\mathcal{O}} \mathfrak{m}_S.$$

[Note: $(R \otimes_{\mathcal{O}} S)/\mathfrak{m} \cong \Bbbk_R \otimes_{\Bbbk_{\mathcal{O}}} \Bbbk_S$ is not necessarily a field, or even a local ring, but it is artinian.] Now define the completed tensor product $R \widehat{\otimes}_{\mathcal{O}} S$ to be the $\mathfrak{m}$-adic completion of $R \otimes_{\mathcal{O}} S$.

Universal property. $R \widehat{\otimes}_{\mathcal{O}} S$ is the coproduct in the category of complete semilocal Noetherian $\mathcal{O}$-algebras and continuous maps. It is thus the universal (i.e. initial) complete semilocal Noetherian $\mathcal{O}$-algebra equipped with continuous $\mathcal{O}$-algebra maps from R and S.

Example. We have $\mathcal{O}[\![x]\!] \widehat{\otimes}_{\mathcal{O}} \mathcal{O}'[\![y]\!] \cong \mathcal{O}'[\![x, y]\!]$ when $\mathcal{O}'$ is any complete Noetherian local $\mathcal{O}$-algebra. We also have

$$(\mathcal{O}[\![x_1, \ldots, x_r]\!]/\mathbb{J}) \widehat{\otimes}_{\mathcal{O}} (\mathcal{O}'[\![y_1, \ldots, y_s]\!]/\mathbb{J}') \cong \mathcal{O}'[\![x_1, \ldots, x_r, y_1, \ldots, y_s]\!]/(\mathbb{J}, \mathbb{J}')$$

in this setup.

4. Global setup and statement

Let F be a number field, and p a prime. Let S be a finite set of places of F containing $\{v|p\}$. Fix an algebraic closure $\overline{F}/F$ and let $F_S \subset \overline{F}$ be the maximal extension unramified outside S. Let $G_{F,S} = \mathrm{Gal}(F_S/F)$. Let $\Sigma \subset S$ be any subset of places [for now; later we'll impose conditions].

For $v \in \Sigma$, fix algebraic closures $\overline{F}_v/F_v$ and choose embeddings $\overline{F} \hookrightarrow \overline{F}_v$, or, what is the same thing, choices of decomposition group $\mathrm{Gal}(\overline{F}_v/F_v) = G_v \subset G_{F,S}$. Now let $K/\mathbb{Q}_p$ be a finite extension, and $\mathcal{O}$, π, and k be as above. Fix a character $\psi : G_{F,S} \to \mathcal{O}^\times$.

Let V_k be a finite dimensional continuous representation of $G_{F,S}$ over k such that $\det V_k = \psi \bmod \pi$.

Since we're fixing $\det = \psi$ in this subsection, we'll be dealing (from now on in this talk) with $\mathrm{ad}^0 V_k$ rather than $\mathrm{ad} V_k$. [More on this later.] A caution is in order: if $p | \dim V_k$ then $\mathrm{ad}^0 V_k$ is not a direct summand of $\mathrm{ad} V_k$. Usually the scalars in $\mathrm{ad} V_k$ give a splitting, but when $p | \dim V_k$ the scalars actually sit inside $\mathrm{ad}^0 V_k$. Hence we shall assume from now on that $p \nmid \dim V_k$.

For each $v \in \Sigma$ fix a basis β_v of V_k. We're going to consider deformation functors (and the representing rings) with determinant conditions. Set $D_v^{\square,\psi}$ to be the functor of framed deformations of $V_k|_{G_v}$ with the basis β_v, with fixed determinant $\psi \bmod \pi$, and let $R_v^{\square,\psi}$ be the ring (pro-)representing it. This always exists. Likewise let $D_{F,S}^{\square,\psi}$ be the functor of deformations V_A of V_k with determinant $\psi \bmod \pi$, equipped with an A-basis $\widetilde{\beta}_v$ of V_A lifting β_v for each $v \in \Sigma$. Let $R_{F,S}^{\square,\psi}$ be the ring representing it. Again, this always exists.

We have analogous respective unframed counterparts R_v^ψ and $R_{F,S}^\psi$ under the usual condition that V_k has only scalar endomorphisms as a representation space for G_v and $G_{F,S}$ respectively.

Now *define* $R_\Sigma^{\square,\psi} = \widehat{\bigotimes}_{v \in \Sigma} R_v^{\square,\psi}$ [completed tensor product over $\mathcal{O}$]. Since each $R_v^{\square,\psi}$ has the same residue field, in this case the completed tensor product actually *is* local! Let $\mathfrak{m}_\psi^\square$ be its maximal ideal. Analogously define R_Σ^ψ and $\mathfrak{m}_\Sigma$. Denote the maximal ideal of the local ring $R_{F,S}^{\square,\psi}$ by $\mathfrak{m}_{F,S}^\square$ and likewise that of $R_{F,S}^\psi$ by $\mathfrak{m}_{F,S}$.

There is a natural R_Σ^ψ-algebra structure on $R_{F,S}^\psi$ via the universal property of $\widehat{\otimes}_{\mathcal{O}}$. Indeed, for each $v \in \Sigma$, by restricting the universal deformation of V_k valued in $R_{F,S}^\psi$ to $G_v \subset G_{F,S}$ the universal property of R_v^ψ induces a canonical local $\mathcal{O}$-algebra morphism $R_v^\psi \to R_{F,S}^\psi$. We then use the universal property of completed tensor products.

Theorem. *For $i \geq 1$ let h_Σ^i (resp. c_Σ^i) denote the k-dimension of the kernel (resp. cokernel) of the map*

$$\theta_i : \mathrm{H}^i(G_{F,S}, \mathrm{ad}^0 V_k) \to \prod_{v \in \Sigma} \mathrm{H}^i(G_v, \mathrm{ad}^0 V_k).$$

Then we have an isomorphism of R_Σ^ψ-algebras

$$R_{F,S}^\psi \cong R_\Sigma^\psi[\![x_1, \ldots, x_r]\!]/(f_1, \ldots, f_{r+s})$$

where $r = h_\Sigma^1$ and $s = c_\Sigma^1 + h_\Sigma^2 - h_\Sigma^1$.

To get the desired presentation, as in the proof of Theorem 1, first consider a surjection

$$\mathbb{B} := R_\Sigma^\psi[\![x_1, \ldots, x_r]\!] \twoheadrightarrow R_{F,S}^\psi$$

where $r = \dim_k \operatorname{coker}(\mathfrak{m}_\Sigma/(\mathfrak{m}_\Sigma^2, \pi) \to \mathfrak{m}_{F,S}/(\mathfrak{m}_{F,S}^2, \pi))$; this surjectivity uses completeness. Dualizing, we have

$$r = \dim_k \ker(\operatorname{Hom}_k(\mathfrak{m}_{F,S}/(\mathfrak{m}_{F,S}^2, \pi), k) \to \operatorname{Hom}_k(\mathfrak{m}_\Sigma/(\mathfrak{m}_\Sigma^2, \pi), k)).$$

Using the computation from Mok's lecture, this is

$$\dim_k \ker \theta_1 = h^1_\Sigma.$$

The key point that makes these computations work is that the completed tensor product represents the product of the functors represented by the R_v^ψ, which is most easily checked by computing on artinian points (for which the completed tensor product collapses to an ordinary tensor product). That then brings us down to the elementary fact that the tangent space of the product of functors is the product of the tangent spaces.

Denote by $\mathbf{m}$ the maximal ideal of $\mathbb{B}$, and by $\mathbb{J}$ the kernel $\ker(\mathbb{B} \twoheadrightarrow R_{F,S}^\psi)$. Now comes a delicate technical point. Like in the proof of Theorem 1, we can set-theoretically lift $\rho : G_{F,S} \to \mathrm{GL}_n(R_{F,S}^\psi)$ to $\widetilde{\rho} : G_{F,S} \to \mathrm{GL}_n(\mathbb{B}/\mathbf{m}\mathbb{J})$, not necessarily a homomorphism, and there arises the problem of finding a continuous such $\widetilde{\rho}$. We seek a better method than the trick as earlier with finite residue fields because we wish to later apply the same technique to future situations involving characteristic-0 deformation theory, for which the residue field is a p-adic field and not a finite field. The reader who prefers to ignore this problem should skip the next section.

5. Continuity nonsense

To explain the difficulty and its solution, let us first formulate a general situation. Consider a surjective map $R' \twoheadrightarrow R$ between complete local noetherian rings with kernel $\mathbb{J}$ killed by $\mathfrak{m}_{R'}$, and assume that we are in one of two cases:

Case 1: residue field k is finite of characteristic p, so R and R' are given the usual max-adic topologies that are profinite. These topologies are the inverse limits of the discrete topologies on artinian quotients.

Case 2 (to come up later!): residue field k is a p-adic field and R and R' are $\mathbb{Q}_p$-algebras, whence uniquely k-algebras in a compatible way (by Hensel). Their artinian quotients are then finite-dimensional as k-vector spaces, and so are naturally topologized as such (making them topological k-algebras, with transition maps that are quotient maps, as for any k-linear surjections between k-vector spaces of finite dimension). Give R and R' the inverse limit of those topologies (which induce the natural k-linear topologies back on the finite-dimensional artinian quotients).

In both cases, let $\rho : G \to \mathrm{GL}_n(R)$ be a continuous representation. We seek to make an obstruction class in a "continuous" $\mathrm{H}^2(G, \operatorname{ad}\rho)$ (over k) for measuring whether or not ρ can be lifted to a *continuous* representation into $\mathrm{GL}_n(R')$. The problem is to determine if ρ has a continuous set-theoretic lifting (moreover with with a fixed determinant if we wish to study deformations with a fixed determinant, assuming that p doesn't divide n).

We saw earlier how to handle Case 1 when R is *artinian*, by a trick. That trick rested on ρ at artinian level factoring through a *finite* quotient of G. Such an argument has no chance of applying when k is a p-adic field in interesting cases, and we're sure going to need that later when studying generic fibers of deformation rings and proving smoothness by proving vanishing of a p-adic H^2. So we need an improvement of the method from artinian Case 1 which addresses the following two points:

(i) what to do when k is p-adic,

(ii) how to incorporate additional things like working with a fixed determinant.

Actually, (ii) will be very simple once we see how to deal with (i), as we will see below. This is important because in practice we want to deal with more general constraints than just "fixed determinant" and so we want a general method which works for any "reasonable property", not just something ad hoc for the property of fixed determinant.

To deal with (i) (and along the way, (ii)), we will use a variant on fix from artinian Case 1. That argument allows us to reduce to deal with the case when R and R' are artinian, but we need to show in that *artinian setting* we can make a continuous set-theoretic lifting without the crutch of "factoring through finite quotient of G" (which is available for finite k but not p-adic k).

First conjugate so the reduction $\rho_0 : G \to \mathrm{GL}_n(k)$ lands in $\mathrm{GL}_n(\mathcal{O}_k)$. Then by using the method from Brian's talk on p-adic points of deformation rings, we can find a finite flat local $\mathcal{O}_k$-algebra $\mathcal{O}_k$-lattice A

inside of R with residue field equal to that of $\mathcal{O}_k$ and containing the compact $\rho(G)$, and then we can find a similar such A' in R' mapping onto A. We'd like to lift

$$\rho : G \to \mathrm{GL}_n(A)$$

to $\mathrm{GL}_n(A')$ set-theoretically in a continuous way. Note that $\mathrm{GL}_n(A') \to \mathrm{GL}_n(A)$ is surjective.

The point is that $\mathrm{GL}_n(A)$ and $\mathrm{GL}_n(A')$ are respectively open in $\mathrm{GL}_n(R)$ and $\mathrm{GL}_n(R')$ with subspace topologies that arise from the ones on A inside R and A' inside R' which are their natural topologies as finite free $\mathcal{O}_k$-modules. This makes them *profinite*, much as $\mathrm{GL}_n(R)$ and $\mathrm{GL}_n(R')$ were in the case of finite k. So we have reduced ourselves to the following situation, in which we will use an argument suggested by Lurie that also gives another approach for handling the case of finite k as well.

Let $H' \to H$ be a continuous surjective map of profinite groups, and $\rho : G \to H$ a continuous homomorphism. We claim that there is a continuous set-theoretic lifting $G \to H'$ of ρ that also respects properties like "fixed det" in the case of intended applications. To see this, let $F \twoheadrightarrow G$ be a surjection from a "free profinite group". The composite map

$$F \twoheadrightarrow G \to H$$

can be lifted continuously to $F \to H'$ even as a homomorphism by individually lifting from H to H' the images of each member of the "generating set" for the free profinite F. Those individual lifts can be rigged to have a desired det, or whatever other "reasonable homomorphic property" can be checked pointwise through a surjection, and so such a property is inherited by the map $F \to H'$. But what about $G \to H'$? If we can find a continuous set-theoretic section of $F \twoheadrightarrow G$ then composing that section with $F \to H'$ will give the required $G \to H'$. So our continuity problems will be settled once we prove the following fact.

Claim: If $f : G' \to G$ is a continuous homomorphism between profinite groups then it has a continuous section (as topological spaces).

Proof. For closed normal subgroups $N' \lhd G'$ and $N := f(N') =$ closed normal in G, consider continuous sections $s : G/N \to G'/N'$ to the induced quotient map $G'/N' \to G/N$ arising from f. For example, such an s exists if $N' = G'$ (so $N = G$). If (N', s) and (M', t) are two such pairs with N' containing M', say $(M', t) \geq (N', s)$ if

$$t : G/M \to G'/M' \qquad \textit{and} \qquad s : G/N \to G'/N'$$

are compatible via the projections $G/M \twoheadrightarrow G/N$ and $G'/M' \twoheadrightarrow G'/N'$.

I claim that the criterion for Zorn's Lemma is satisfied. Let $\{(N'_i, s_i)\}$ be a chain of such pairs, and let $N' = \bigcap N'_i$. Then the natural map

$$G'/N' \to \varprojlim G'/N'_i$$

is surjective (since an inverse limit of surjections $G'/N' \to G'/N'_i$ between compact Hausdorff spaces), yet also injective and thus a homeomorphism. Likewise, for $N := \bigcap N_i$ the map $G/N \to \varprojlim G/N_i$ a homeomorphism, and I claim that $N = f(N')$. Indeed, if x is in N then $f^{-1}(x)$ meets each N'_i in a non-empty closed set, and these satisfy the finite intersection property since $\{N'_i\}$ is a chain ordered by inclusion, so $f^{-1}(x)$ contains a point in the intersection N' of all N'_i. That says x is in $f(N')$ as desired. (The inclusion of $f(N')$ inside of N is clear.)

It follows that the compatible continuous sections $s_i : G_i/N_i \to G'_i/N'_i$ induced upon passing to the projective limit define a continuous section

$$s : G/N \to G'/N',$$

so (N', s') is an upper bound on the chain $\{(N'_i, s_i)\}$.

Now we apply Zorn's Lemma to get a maximal element (N', s). This is a continuous section $s : G/N \to G'/N'$ where $N = f(N')$. I claim $N' = \{1\}$, so we will be done. If not, then since $N' \cap U'$ for open normal subgroups U' in G' define a base of opens in N' around 1 (as N' gets its profinite topology as subspace topology from G'), there must exist such U' so that $N' \cap U'$ is a proper subgroup of N'. Replacing G' with $G'/(N' \cap U')$ and G with quotient by image of $N' \cap U'$ in G brings us to the case where N is finite and *non-trivial* yet (N', s) retains the maximality property (no continuous section using a proper [closed] subgroup of N' normal in G'). We seek a contradiction.

Since N' and N are finite, the quotient maps $q' : G' \twoheadrightarrow G'/N'$ and $q : G \twoheadrightarrow G/N$ are covering spaces with finite constant degree > 0. By total disconnectedness, these covering spaces admit sections. Composing s

with a section to q' gives a continuous section $G/N \to G'$ to

$$G' \xrightarrow{f} G \xrightarrow{q} G/N.$$

Composing such a section with q gives a continuous map $t : G \to G'$ so that $f(t(g)) = g \bmod N$, so by profiniteness of G and finiteness of N we get an open normal subgroup U in G such that for each representative g_i of G/U there exists $n_i \in N$ such that $f(t(g_i u)) = n_i g_i u$ for all $u \in U$. But $n_i = f(n_i')$, so replacing t on $g_i U$ with $(n_i')^{-1} t$ for each i gives a new t so that $f(t(g_i u)) = g_i u$ for all $u \in U$ and all i, which is to say $ft = \mathbf{1}_G$. This exhibits a continuous section t to f, contradicting that N was arranged to be nontrivial and maximal with respect to the preceding Zorn's Lemma construction. Hence, in fact N above is $\{1\}$ so we are done. □

6. Proof of Theorem 4

Returning to the situation of interest, we now have a continuous $\widetilde{\rho}$ that can even be arranged to satisfy $\det \widetilde{\rho} \equiv \psi \bmod \mathbf{m}\mathbb{J}$. Still following the argument from the proof of Theorem 1, define for $f \in \mathrm{Hom}_k(\mathbb{J}/\mathbf{m}\mathbb{J}, k)$ the continuous 2-cocycle φ_f as before, and observe that this time the determinant condition entails that $[\varphi_f] \in \mathrm{H}^2(G_{F,S}, \mathrm{ad}^0 V_k)$. The proof of the well-definedness of $[\varphi_f]$ is as before. Also we still have the equivalence that $[\varphi_f] = 0$ if and only if $\widetilde{\rho}$ can be chosen to be a homomorphism $\bmod \ker f$.

Now for the restriction of ρ to each G_v, we know we can find a continuous lift, namely coming from the universal representation ρ_v at v:

$$G_v \xrightarrow{\rho_v} \mathrm{GL}_n(R_v^\psi) \to \mathrm{GL}_n(R_\Sigma^\psi) \to \mathrm{GL}_n(\mathbb{B})$$

where the other maps are the obvious ones. Hence the class $[\varphi_f]|_{G_v} \in \mathrm{H}^2(G_v, \mathrm{ad}^0 V_k)$ is always trivial. In other words, we have a k-linear map $\mathrm{Hom}_k(\mathbb{J}/\mathbf{m}\mathbb{J}, k) \xrightarrow{\Phi} \ker \theta_2$ satisfying $f \mapsto [\varphi_f]$; the target has dimension h^2_Σ by definition. Therefore [easy exercise] it suffices to show that $\dim_k \ker \Phi \le c^1_\Sigma$. (All we need is the inequality, because we can always throw in extra trivial "relations" $f_i = 0$ into the denominator of $R^\psi_{F,S}$.)

Let $I = \ker(\mathfrak{m}_\Sigma/(\mathfrak{m}_\Sigma^2, \pi) \to \mathfrak{m}_{F,S}/(\mathfrak{m}_{F,S}^2, \pi))$. Then $\mathrm{Hom}_k(I, k) \cong \mathrm{coker}(\theta_1)$. So it is enough to construct a linear injection $\ker \Phi \hookrightarrow \mathrm{Hom}_k(I, k)$.

Step 1: Observe that $I = \ker(\mathbf{m}/(\mathbf{m}^2, \pi) \to \mathfrak{m}_{F,S}/(\mathfrak{m}_{F,S}^2, \pi))$ because we chose the x_i's to map onto a basis of $\mathrm{coker}(\mathfrak{m}_\Sigma/(\mathfrak{m}_\Sigma^2, \pi) \to \mathfrak{m}_{F,S}/(\mathfrak{m}_{F,S}^2, \pi))$. (In other words, none of the extra stuff in $\mathbf{m}$ dies when we map to $\mathfrak{m}_{F,S}$.)

Step 2: We next claim that $\mathbb{J}/\mathbf{m}\mathbb{J}$ surjects onto I. To prove this, first note that the map $\mathbb{J}/\mathbf{m}\mathbb{J} \to \mathbf{m}/(\mathbf{m}^2, \pi)$ comes from tensoring

$$0 \to \mathbb{J} \to \mathbf{m} \to \mathfrak{m}_{F,S} \to 0$$

over $\mathbb{B}$ with $\mathbb{B}/\mathbf{m}$ and then reducing mod π. We need to show that this map is surjective onto I. Fix $x \in I \subset \mathbf{m}/(\mathbf{m}^2, \pi)$. We know

$$\mathbb{J}/\mathbf{m}\mathbb{J} \twoheadrightarrow \ker(\mathbf{m}/\mathbf{m}^2 \to \mathfrak{m}_{F,S}/\mathfrak{m}_{F,S}^2).$$

We can lift x to $\widetilde{x} \in \mathbf{m}/\mathbf{m}^2$. Since x maps to zero in $\mathfrak{m}_{F,S}/(\mathfrak{m}_{F,S}^2, \pi)$, $\widetilde{x}$ maps to $\pi r \bmod \mathfrak{m}_{F,S}^2$ for some $r \in R^\psi_{F,S}$. But now we can just choose some $\widetilde{r} \in \mathbb{B}$ mapping to $r \in R^\psi_{F,S}$ (i.e. mod $\mathbb{J}$). Now replace $\widetilde{x}$ with $\widetilde{x} - (\pi\widetilde{r} \bmod \mathfrak{m}^2)$ so that $\widetilde{x}$ has vanishing image in $\mathfrak{m}_{F,S}/\mathfrak{m}_{F,S}^2$. That says $\widetilde{x}$ is in the image of $\mathbb{J}/\mathfrak{m}\mathbb{J}$ in $\mathfrak{m}/\mathfrak{m}^2$, so x is hit by $\mathbb{J}/\mathfrak{m}\mathbb{J}$ as desired.

Step 3: By Step 2 we get $\mathrm{Hom}_k(I, k) \hookrightarrow \mathrm{Hom}_k(\mathbb{J}/\mathbf{m}\mathbb{J}, k) \supset \ker \Phi$. So we need to show that $\ker \Phi \subset \mathrm{Hom}_k(I, k)$. In other words, if $[\varphi_f] = 0$ then we claim that $f : \mathbb{J}/\mathbf{m}\mathbb{J} \to k$ should factor through I, or equivalently vanish on $K = \ker(\mathbb{J}/\mathbf{m}\mathbb{J} \twoheadrightarrow I)$. Or equivalently, we need to show that $K = \mathbb{J} \cap (\mathbf{m}^2, \pi) \subset \mathbb{J}_f = \ker f$. But in fact this is really what we showed at the end of the proof of Theorem 1 when we showed property (4) of Φ.

7. The framed case

Let

$$\eta : \mathfrak{m}_\Sigma^\square/((\mathfrak{m}_\Sigma^\square)^2, \pi) \to \mathfrak{m}_{F,S}^\square/((\mathfrak{m}_{F,S}^\square)^2, \pi).$$

Then

$$R^{\square,\psi}_{F,S} \cong R^{\square,\psi}_\Sigma[\![x_1, \ldots, x_{r^\square}]\!]/(f_1, \ldots, f_{r^\square + s^\square}),$$

where $r^\square = \dim_k \operatorname{coker}\eta$ and $r^\square + s^\square = h^2_\Sigma + \dim_k \ker\eta$.

The proof is the same as in the unframed case, just with extra squares floating around all over the place. But now our H's have turned into Z's (that is, elements of the tangent space which were cohomology classes are now cocycles) so it's better to phrase the result as above.

8. Formulas for r's and s's

Theorem. *Suppose that $\{v|p\} \subset \Sigma$, that $\{v|\infty\} \subset S$, and that $S - \Sigma$ contains at least one finite prime. Then (with notation as above)*

$$s = \sum_{v|\infty, v\notin\Sigma} \dim_k (\mathrm{ad}^0 V_k)^{G_v}.$$

Remark. We also have $r^\square \geq \#\Sigma - 1, r^\square \stackrel{?}{=} r + \#\Sigma - 1, s^\square = s - \#\Sigma + 1$.

Proof. Let $Y = \mathrm{ad}^0 V_k$ and $X = Y^\vee(1)$. (In the notation of Rebecca's talk, $X = Y'$; it is written as a "twisted Pontrjagin dual" here because instead of being Hom into $\mathbb{Q}/\mathbb{Z}$ (trivial G-module) the target is given the action of the cyclotomic character.) Recall the end of the Poitou-Tate exact sequence (from Rebecca's talk)

$$\mathrm{H}^2(G_{F,S}, Y) \to \prod_{v\in S} \mathrm{H}^2(G_v, Y) \to \mathrm{H}^0(G_{F,S}, X)^\vee \to 0.$$

Split the product into two pieces:

$$\prod_{v\in S} \mathrm{H}^2(G_v, Y) = \prod_{v\in\Sigma} \mathrm{H}^2(G_v, Y) \times \prod_{v\in S-\Sigma} \mathrm{H}^2(G_v, Y).$$

The claim is that as long as the second factor is nonzero (which it is by hypothesis), it surjects onto $\mathrm{H}^0(G_{F,S}, X)^\vee$. Indeed, trivially $\mathrm{H}^0(G_{F,S}, X) \hookrightarrow \mathrm{H}^0(G_v, X)$ since restricting to the decomposition group gives more invariants. Dually, we have $\mathrm{H}^0(G_v, X)^\vee \twoheadrightarrow \mathrm{H}^0(G_{F,S}, X)^\vee$. But by the Tate pairing, $\mathrm{H}^0(G_v, X)^\vee \cong \mathrm{H}^2(G_v, Y)$. On each factor, the last map in the Tate-Poitou sequence is none other than the composition $\mathrm{H}^2(G_v, Y) \cong \mathrm{H}^0(G_v, X)^\vee \twoheadrightarrow \mathrm{H}^0(G_{F,S}, X)^\vee$. Thus the claim is true.

Now we do a little diagram chase. We have

$$\mathrm{H}^2(G_{F,S}, Y) \to \prod_{v\in\Sigma} \mathrm{H}^2(G_v, Y) \times \prod_{v\in S-\Sigma} \mathrm{H}^2(G_v, Y) \to \mathrm{H}^2(G_{F,S}, X)^\vee \to 0.$$

The claim is that $\mathrm{H}^2(G_{F,S}, Y) \twoheadrightarrow \prod_{v\in\Sigma} \mathrm{H}^2(G_v, Y)$. Indeed, given $(a_v)_\Sigma \in \prod_{v\in\Sigma} \mathrm{H}^2(G_v, Y)$, suppose its image in $\mathrm{H}^2(G_{F,S}, X)^\vee$ is γ. Since $\prod_{v\in S-\Sigma} \mathrm{H}^2(G_v, Y) \twoheadrightarrow \mathrm{H}^2(G_{F,S}, X)^\vee$, we can find

$$(b_v)_{S-\Sigma} \in \prod_{v\in S-\Sigma} \mathrm{H}^2(G_v, Y)$$

such that the image of $(b_v)_{S-\Sigma}$ in $\mathrm{H}^2(G_{F,S}, X)^\vee$ is $-\gamma$. Then

$$(a_v)_\Sigma \times (b_v)_{S-\Sigma} \in \ker(\prod_S \mathrm{H}^2(G_v, Y) \twoheadrightarrow \mathrm{H}^2(G_{F,S}, X)^\vee),$$

whence this tuple is in the image of $\mathrm{H}^2(G_{F,S}, Y)$. Projecting onto the $\prod_{v\in\Sigma}$ factor proves the claim. But the surjectivity of $\mathrm{H}^2(G_{F,S}, Y) \twoheadrightarrow \prod_{v\in\Sigma} \mathrm{H}^2(G_v, Y)$ says precisely that $c^2_\Sigma = \dim\operatorname{coker}\theta_2 = 0$.

Consequently we have $h^2_\Sigma = h^2(G_{F,S}, Y) - \sum_{v\in\Sigma} h^2(G_v, Y)$. So by the formulas at the end of Theorem 4,

$$s = -h^1_\Sigma + c^1_\Sigma + h^2_\Sigma = -h^1(G_{F,S}, Y) + \sum_{v\in\Sigma} h^2(G_v, Y) + h^2(G_{F,S}, Y) - \sum_{v\in\Sigma} h^2(G_v, Y).$$

Now recall that we have assumed throughout that $\mathrm{End}_{G_{F,S}} V_k = (\mathrm{ad}\, V_k)^{G_{F,S}} = k$ (since we need this to make sure the unframed deformation ring even exists!). In particular, $(\mathrm{ad}^0 V_k)^{G_{F,S}} = 0$. That is, $h^0(G_{F,S}, Y) = h^0(G_v, Y) = 0$. So we can add $h^0(G_{F,S}, Y) - \sum_{v\in\Sigma} h^0(G_v, Y)$ to s and nothing changes. But now we recognize from the equation above that in fact $s = \chi(G_{F,S}, Y) - \sum_{v\in\Sigma} \chi(G_v, Y)$.

We now invoke the Tate global Euler characteristic formula. [Reference: Milne, *Arithmetic Duality Theorems* Ch. I, Thm. 5.1.] We conclude that

$$\chi(G_{F,S}, Y) = \sum_{v|\infty} h^0(G_v, Y) - [F:\mathbb{Q}] \dim_k Y.$$

We also have for $v < \infty, v \nmid p$, that $\chi(G_v, Y) = 0$. For $v < \infty, v|p$, we have $\chi(G_v, Y) = -[F_v : \mathbb{Q}_p] \dim_k Y$. For $v|\infty$, we have $\chi(G_v, Y) = h^0(G_v, Y)$. One sees that in $s = \chi(G_{F,S}, Y) - \sum_{v \in \Sigma} \chi(G_v, Y)$, the degree contributions all cancel out, so there are no non-archimedean terms. Of the archimedean places, all those in Σ cancel as well, and we are left with the statement of the theorem. $\square$

Notes on Galois Cohomology—Modularity seminar

Rebecca Bellovin

1 Introduction

We've seen that the tangent space to a deformation functor is a Galois cohomology group H^1, and we'll see that obstructions to a deformation problem will be in H^2. So if we want to know things like the dimension of R or whether a deformation functor is smooth, we need to be able to get our hands on the cohomology groups. Secondarily, if we want to "deform subject to conditions", we'll want to express the tangent space and obstruction space of those functors as cohomology groups, and cohomology groups we can compute in terms of an unrestricted deformation problem.

For the most part, we will assume the contents of Serre's *Local Fields* and *Galois Cohomology*. These cover the cases when G is finite (and discrete) and M is discrete, and G is profinite and M is discrete, respectively.

References:
Serre's *Galois Cohomology*
Neukirch's *Cohomology of Number Fields*
Appendix B of Rubin's *Euler Systems*
Washington's article in CSS
Darmon, Diamond, and Taylor (preprint on Darmon's website)

2 Generalities

Let G be a group, and let M be a module with an action by G. Both G and M have topologies; often both will be discrete (and G will be finite),

or G will be profinite with M discrete; or both will be profinite. We always require the action of G on M to be continuous.

Let's review group cohomology, using inhomogenous cocycles.

For a topological group G and a topological G-module M, the ith group of continuous cochains $C^i(G, M)$ is the group of continuous maps $G^i \to M$. There is a differential $d : C^i(G, M) \to C^{i+1}(G, M)$ given by

$$\begin{aligned}(df)(g_1, \ldots, g_{i+1}) = \; & g_1 \cdot f(g_2, \ldots, g_{i+1}) \\ & + \sum_{j=1}^{n} (-1)^j f(g_1, \ldots, g_j g_{j+1}, \ldots, g_{i+1}) \\ & + (-1)^{i+1} f(g_1, \ldots, g_i)\end{aligned}$$

It is easy to check that d^2 is zero, so we have a complex $C^\bullet(G, M)$. Then we define $H^i(G, M) := \ker d / \operatorname{im} d$.

If G is finite and M is discrete, this is just ordinary group cohomology, see for example [3]. But for G or M profinite, taking the algebraic group cohomology gives the "wrong" answer. For example, if L/K is a finite Galois extension of fields and M is a module equipped with a trivial action of $G := \mathrm{Gal}(L/K)$, then the algebraic cohomology group $H^1(G, M) = \mathrm{Hom}(G, M)$ classifies subextensions $K \subset K' \subset L$ with $\mathrm{Gal}(K'/K)$ isomorphic to a subgroup of M. It would be nice if we could relax the finiteness hypothesis on the extension L/K and still have H^1 meaningfully classify subextensions. But infinite Galois theory tells us that only *closed* subgroups of $\mathrm{Gal}(L/K)$ correspond to subextensions $K \subset K' \subset L$, so our definition of H^1 will have to take topological information into account somehow.

For an explicit example where algebraic and continuous group cohomology differ, see Brian's notes from Hawaii, exercise 2.5.2.

2.1 Functorial properties

As we have defined it, Galois cohomology is functorial in the coefficients, that is, given a morphism $M \to M'$ of G-modules, there is are morphisms $H^i(G, M) \to H^i(G, M')$. Suppose $0 \to M' \to M \to M'' \to 0$ is an exact sequence of topological modules, and there is continuous section $M'' \to M$ (as sets, not modules!). Then $0 \to C^i(G, M') \to C^i(G, M) \to C^i(G, M'') \to 0$ is

exact for every i, and by homological algebra nonsense, we get a long exact sequence

$$\cdots \to H^i(G, M') \to H^i(G, M) \to H^i(G, M'') \to H^{i+1}(G, M') \to \cdots$$

In all the cases we will care about, this hypothesis will be satisfied, because surjective maps of discrete topological spaces have continuous sections, and proposition 1, chapter 1 of *Galois Cohomology* tells us that continuous surjections of profinite groups have continuous sections. In particular, if M is a finitely-generated $\mathbb{Z}_p$-module or a finite-dimensional $\mathbb{Q}_p$-vector space, we will have a long exact sequence.

For finite groups G and discrete G-modules M, recall that for all subgroups $H \subset G$, we have a restriction map

$$\text{res} : H^i(G, M) \to H^i(H, M)$$

and a corestriction map

$$\text{cor} : H^i(H, M) \to H^i(G, M)$$

If H is normal in G, we also have an inflation map

$$\text{inf} : H^i(G/H, M^H) \to H^i(G, M)$$

For G profinite and M discrete, we still have a restriction map

$$\text{res} : H^i(G, M) \to H^i(H, M)$$

If H is a closed, normal subgroup of G (so that the quotient G/H makes sense), we also still have an inflation map

$$\text{inf} : H^i(G/H, M^H) \to H^i(G, M)$$

However, to define a corestriction map, we need to assume H is open in G with finite index. In that case, we define it "at finite level" (as discussed in section 2.2) using the definition from finite group cohomology, and take the limit.

When G is a finite group, or G is profinite and M is discrete, for any normal subgroup H there is a spectral sequence $H^p(G/H, H^q(H, M)) \to$

$H^{p+q}(G, M)$. This is because cohomology groups $H^q(G, M)$ are the derived functors (taken in the category of all G modules if G is finite, but taken in the category of discrete G-modules if G is profinite) of the functor $M \mapsto M^G$, and $M \mapsto M^G$ is the composition of $M \mapsto M^H$ and $M^G \mapsto (M^H)^{G/H}$. In particular, the low-degree terms of the spectral sequence give us the Hochschild-Serre exact sequence

$$0 \to H^1(G/H, M^H) \overset{\text{inf}}{\to} H^1(G, M) \overset{\text{res}}{\to} H^1(H, M)^{G/H} \to H^2(G/H, M^H) \overset{\text{inf}}{\to} H^2(G, M)$$

The first four terms are the usual inflation-restriction exact sequence.

Recall also that in finite group cohomology, there is a cup-product pairing $H^p(G, M) \times H^q(G, N) \overset{\cup}{\to} H^{p+q}(G, M \otimes N)$ given on the level of cochains by $(\varphi \cup \psi)(g_1, \ldots, g_p, g_{p+1}, \ldots, g_{p+q}) = \varphi(g_1, \ldots, g_p) \otimes g_1 \cdots g_p \psi(g_{p+1}, \ldots, g_{p+q})$. The same applies for profinite groups G and discrete G-modules. However, if we want to allow more interesting topologies on the coefficient modules, we may not be able to define the tensor product of modules. Instead, we use the same formula to say that whenever there are (continuous) maps of G-modules $M \to P, N \to P$, there is a cup-product $H^p(G, M) \times H^1(G, N) \overset{\cup}{\to} H^{p+q}(G, P)$.

2.2 Reducing to the Finite/Discrete Case

Now let's allow G to be profinite (still assuming M to be discrete).

Theorem 2.1. *Let (G_i) be a projective system of profinite groups, and let (M_i) be an inductive system of discrete G_i-modules (the maps are all compatible). If $G = \varprojlim G_i$ and $M = \varinjlim M_i$, then $H^q(G, M) = \varinjlim H^q(G_i, M_i)$.*

In particular,

Corollary 2.2. *For profinite G, $H^q(G, M) = \varinjlim H^q(G/U, M^U)$ for $q \geq 0$, where the limit is taken over all open normal subgroups of G.*

This corollary lets us reduce many statements to the equivalent statements at finite level. For example, classical group cohomology tells us that for a finite group G, $H^q(G, M)$ is torsion for $q \geq 1$, so for profinite G, $H^q(G, M)$ is the colimit of torsion groups, so is itself torsion.

It also lets us make definitions at finite level, and then take a direct limit. For example, in order to define corestriction for profinite groups, we recall the definition of the corestriction map $\mathrm{cor} : H^q(H/(H \cap U), M) \to H^q(G/U, M)$ for open normal subgroups $U \subset G$ of finite index. By applying the above corollary, we obtain a homomorphism $\mathrm{cor} : H^q(H, M) \to H^q(G, M)$.

Now let's relax the assumption that M is discrete. Then we have the following results due to Tate (see [4] or Appendix B of [1]):

Proposition 2.3. *For $T = \varprojlim T_n$, T_n finite, if $i > 0$ and $H^{i-1}(G, T_n)$ is finite for every n, then $H^i(G, T) = \varprojlim H^i(G, T_n)$.*

Proposition 2.4. *If T is a finitely generated $\mathbb{Z}_p$-module and $i \geq 0$, then $H^i(G, T)$ has no divisible elements, and $H^i(G, T) \otimes \mathbb{Q}_p \tilde{\to} H^i(G, T \otimes \mathbb{Q}_p)$.*

If we wanted, we could have first defined group cohomology for discrete G-modules, and then defined $H^i(G, T)$ by $\varprojlim H^i(G, T_n)$ and $H^i(G, T \otimes \mathbb{Q}_p)$ by $H^i(G, T) \otimes \mathbb{Q}_p$, instead of via continuous cochains. Then these propositions show we would end up with the same theory (at least for the coefficient modules we care about).

These propositions also give us generalizations of the inflation-restriction exact sequence and the five-term exact sequence associated to the Hochschild-Serre spectral sequence.

Proposition 2.5. *Suppose H is a closed normal subgroup of G.*

1. *There is an inflation-restriction exact sequence*

$$0 \to H^1(G/H, T^H) \to H^1(G, T) \to H^1(H, T)$$

2. *Suppose that p is a prime and for every G-module (resp. H-module) N of finite p-power order, $H^1(G, N)$ and $H^2(G, N)$ (resp. $H^1(H, N)$) is finite. If M is discrete or a finitely generated $\mathbb{Z}_p$-module or a finite-dimensional $\mathbb{Q}_p$-vector space, then there is a Hochschild-Serre exact sequence*

$$0 \to H^1(G/H, M^H) \to H^1(G, M) \to H^1(H, M)^{G/H} \to H^2(G/H, M^H) \to H^2(G, T)$$

2.3 New Phenomena

However, there are some genuinely new phenomena when our groups are profinite, even if our coefficients are still discrete. For example, there is the notion of cohomological dimension:

Definition 2.6. *Let p be a prime and G a profinite group. If for every discrete torsion G-module M and for every $q > n$, the p-primary component of $H^q(G, M)$ is zero, and n is the smallest integer with these properties, we say that n is the p-cohomological dimension of G and denote it by* $\mathrm{cd}_p(G)$.

Removing the requirement that the coefficients be torsion, we make the following definition:

Definition 2.7. *Let p be a prime and G a profinite group. If for every discrete G-module M and for every $q > n$, the p-primary component of $H^q(G, M)$ is zero, and n is the smallest integer with this property, we say that n is the strict p-cohomological dimension of G and denote it by* $\mathrm{scd}_p(G)$.

Of course, we could have infinite cohomological dimension or strict cohomological dimension.

Note that these are not interesting concepts when G is assumed finite! Recall that for any finite cyclic group G, $H^0_T(G, \mathbb{Z}) = \mathbb{Z}/\#G\mathbb{Z}$ and $H^r_T(G, \mathbb{Z}) \cong H^{r+2}_T(G, \mathbb{Z})$ for all $r \in \mathbb{Z}$.

Examples:

- Let $G = \hat{Z}$. Then for every p, $\mathrm{cd}_p(G) = 1$ (see [3, Ch. XIII, Prop. 2]). But $H^2(G, \mathbb{Z}) \cong H^1(G, \mathbb{Q}/\mathbb{Z}) = \mathbb{Q}/\mathbb{Z}$, so $\mathrm{scd}_p(G) = 2$.
- Let G_ℓ be the absolute Galois group of $\mathbb{Q}_\ell$. Then for all p, $\mathrm{cd}_p(G_\ell) = \mathrm{scd}_p(G_\ell) = 2$. This is a manifestation of the general fact that if k is the residue field of K, then $\mathrm{cd}_p(G_K) \leq 1 + \mathrm{cd}_p(G_k)$, with equality when $\mathrm{cd}_p(G_k) < \infty$ and p is different than the characteristic.

3 Local Duality

Now let's try to say something about group cohomology we care about as number theorists. Let K be a p-adic field, i.e., a finite extension of $\mathbb{Q}_p$ and let μ_n be the group of nth roots of unity in $\overline{K}$.

From now on, we will be considering Galois cohomology, that is, group cohomology where the groups in question are Galois groups. If K is a field, we will write $H^i(K, M)$ to mean $H^i(G_K, M)$, and if K'/K is a Galois extension of fields, we will write $H^i(K'/K, M)$ to mean $H^i(\mathrm{Gal}(K'/K), M)$.

Proposition 3.1.
- $H^0(K, \mu_n) = \mu_n \cap K$
- $H^1(K, \mu_n) = K^\times/(K^\times)^n$
- $H^2(K, \mu_n) = \mathbb{Z}/n\mathbb{Z}$
- $H^i(K, \mu_n) = 0$ *for* $i \geq 3$

Proof. The first assertion follows by definition. For the cases $i = 1$ and $i = 2$, use the exact sequence $0 \to \mu_n \to \mathbb{G}_m \to \mathbb{G}_m \to 0$ and look at the long exact sequence in cohomology:

$$\begin{aligned} 0 \quad &\to \quad H^0(K, \mu_n) \to H^0(K, \mathbb{G}_m) \xrightarrow{n} H^0(K, \mathbb{G}_m) \to \\ &\to \quad H^1(K, \mu_n) \to H^1(K, \mathbb{G}_m) \xrightarrow{n} H^1(K, \mathbb{G}_m) \to \\ &\to \quad H^2(K, \mu_n) \to H^2(K, \mathbb{G}_m) \xrightarrow{n} H^2(K, \mathbb{G}_m) \end{aligned}$$

By Hilbert's Satz 90, $H^1(K, \mathbb{G}_m) = 0$, which implies that $H^1(K, \mu_n) = K^\times/(K^\times)^n$. In addition, $H^2(K, \mathbb{G}_m) = \mathbb{Q}/\mathbb{Z}$, with the isomorphism given by the inv map, by the theory of Brauer groups. This implies $H^2(K, \mu_n) = \mathbb{Z}/n\mathbb{Z}$. For $i \geq 3$, the assertion is a theorem of Tate, and is proved in ([2, §4.3, Prop. 12]). □

In particular, this has the striking corollary

Corollary 3.2. *For M a finite G_K-module, $H^i(K, M)$ is finite as well.*

Proof. Over a finite extension K'/K, M becomes a $G_{K'}$-module isomorphic to a direct sum of μ_n's. We have a spectral sequence $H^i(\mathrm{Gal}(K'/K), H^j(K', M)) \Rightarrow H^{i+j}(K, M)$, so by the proposition, $H^{i+j}(K, M)$ is finite. □

Now we can state Tate's local duality theorem:

Theorem 3.3. *Let M be a finite G_K-module and set $M' = \mathrm{Hom}(M, \mu) = \mathrm{Hom}(M, \mathbb{G}_m)$. Then for $0 \leq i \leq 2$, the cup-product*

$$H^i(K, M) \times H^{2-i}(K, M') \to H^2(K, \mu) = \mathbb{Q}/\mathbb{Z}$$

is a perfect pairing.

One of the consequences is the Euler-Poincaré characteristic. For a finite G_K-module M, we define the Euler-Poincaré characteristic to be

$$\chi(M) := \frac{\#H^0(K,M)\#H^2(K,M)}{\#H^1(K,M)}$$

Then one can show that $\chi(M) = p^{-v_p(\#M)\cdot N} = 1/(\mathcal{O} : \#M\mathcal{O})$, where $N = [K : \mathbb{Q}_p]$ and $\mathcal{O}$ is the ring of integers of K. In particular, if the order of A is relatively prime to p, then $\chi(A) = 1$.

We can extend the concept to the case where M is a finite free $\mathbb{Z}_\ell$-module or a finite-dimensional $\mathbb{Q}_\ell$-vector space by making the more familiar definition

$$\chi(M) := h^0(M) - h^1(M) + h^2(M)$$

where $h^i(M) := \operatorname{rk} H^i(K,M)$. If M is a free $\mathbb{Z}_\ell$-module of rank k, take $M_n = M/\ell^n M$, so that $\chi(M) = \varprojlim \frac{1}{n}\log_\ell \chi(M_n) = -kNv_p(\ell)$. In particular, if $\ell \neq p$, then $\chi(M) = 0$.

Here are some interesting special cases:

- Take $M = \mathbb{Z}/n\mathbb{Z}$ and $i = 1$. Then this theorem says we have a perfect pairing $H^1(K, \mathbb{Z}/n\mathbb{Z}) \times H^1(K, \mu_n) \to \mathbb{Q}/\mathbb{Z}$, which in particular says that $\operatorname{Hom}(G_K, \mathbb{Z}/n\mathbb{Z})$ is dual to $K^\times/(K^\times)^n$. This is the duality given by local class field theory, and if K contains the nth roots of unity, Tate duality becomes the Hilbert symbol $K^\times/(K^\times)^n \times K^\times/(K^\times)^n \to \mathbb{Q}/\mathbb{Z}$.

- If E is an elliptic curve (or A is an abelian variety) over K, there is an action of G_K on the torsion $E(\overline{K})[m]$, so we have the perfect pairing

  $$H^1(K, E(\overline{K})[m]) \times H^1(K, E(\overline{K})[m]') \to \mathbb{Q}/\mathbb{Z}$$

  But the Weil pairing tells us that $E(\overline{K})[m]$ is dual to $\hat{E}(\overline{K})[m]$, which for elliptic curves implies we have a pairing

  $$H^1(K, E(\overline{K})[m]) \times H^1(K, E(\overline{K})[m]) \to \mathbb{Q}/\mathbb{Z}$$

3.1 Unramified Cohomology

We're going to be interested in a subgroup of H^1 called the unramified cohomology. We define

$$H^i_{nr}(K,M) := H^i(K^{nr}/K, M^I)$$

to be the cohomology classes vanishing on inertia. For example,

- $H^0_{nr}(K,M) = H^0(K,M)$
- $H^1_{nr}(K,M) = \ker(H^1(K,M) \to H^1(K^{nr},M))$—this is by inflation-restriction. If M is finite, the order of $H^1_{nr}(K,M)$ is the same as the order of $H^0(K,M)$, because there is an exact sequence

  $$0 \to M^{G_K} \to M^I \stackrel{\mathrm{Frob}-\mathrm{id}}{\to} M^I/(\mathrm{Frob}-\mathrm{id})M^I \to 0$$

  The lefthand term is $H^0(K,M)$ and the righthand term is $H^1_{nr}(K,M)$.
- $H^i_{nr}(K,M) = 0$ for $i \geq 2$ because $\mathrm{Gal}(K^{nr}/K) = \hat{\mathbb{Z}}$ has cohomological dimension 1.

Why do we care? For one thing, suppose ρ is an unramified representation and $c \in H^1_{nr}(K,M)$, and consider the corresponding deformation ρ'. Then ρ' restricted to I is the trivial deformation, so ρ' is still unramified.

Going back to elliptic curves, let's briefly make K a global field with E an elliptic curve (or abelian variety) defined over it. Define the finite set of places S to be the union of the archimedean places, the places where E has bad reduction, and the places v where $v(m) \neq 0$, and define K_S to the be the maximal extension of K unramified outside of S. Then $E[m]$ is a G_K-module, so we have the exact sequence

$$0 \to E(\overline{K})[m] \to E(\overline{K}) \stackrel{m}{\to} E(\overline{K}) \to 0$$

The long exact sequence in cohomology gives us

$$0 \to E(K)[m] \to E(K) \stackrel{m}{\to} E(K) \to H^1(G_K, E(\overline{K})[m])$$

so

$$0 \to E(K)/mE(K) \to H^1(G_K, E(\overline{K})[m])$$

We are interested in $E(K)/mE(K)$ because of its role in the proof of the Mordell-Weil theorem. In fact, its image in $H^1(K_S, E(\overline{K})[m])$ is exactly the subgroup of cohomology classes unramified outside S.

Going back to the general theory, let's look at what happens in the Tate pairing. I claim that if $\#M$ is relatively prime to p, then $H^1_{nr}(K,M)$ and

$H^1_{nr}(K, M')$ exactly annhilate each other. To see this, note that the inclusion $H^1_{nr}(K, M) \hookrightarrow H^1(K, M)$ is compatible with cup-product, so the cup-product map

$$H^1_{nr}(K, M) \times H^1_{nr}(K, M') \to H^2(K, \overline{K}^\times)$$

factors through $H^2_{ur}(K, \overline{K}^\times)$, which is zero. So we only need to check that the orders of $H^1_{nr}(K, M)$ and $H^1_{nr}(K, M')$ match up, i.e., that $\#H^1(K, M)/\#H^1_{nr}(K, M) = \#H^1_{nr}(K, M')$. By the argument above, $H^1_{nr}(K, M)$ has the same number of elements as $H^0(K, M)$, and $H^1_{nr}(K, M')$ has the same number of elements as $H^0(K, M')$, which is identified with $H^2(K, M)$ by Tate duality. Since $\#M$ is relatively prime to p, the Euler characteristic of M is 1, which implies the desired equality.

4 Global Euler Characteristic and Poitou-Tate Long Exact Sequence

4.1 Local Conditions

We are going to care about deformation problems more restricted than "all deformations to A", and we'll want to identify tangent spaces of restricted problems with cohomology groups, ideally subgroups of the cohomological tangent spaces we already know about. For example, if we ask for deformations preserving the determinant, we find that the tangent space is $H^1(G, \mathrm{ad}^0 \rho)$: let $C : G \to \mathrm{ad}\,\rho$ be the cocycle representing an infinitesimal deformation, i.e., the deformation is $\rho'(g) = (I + \varepsilon C(g))\rho(g)$. Then $\det(\rho') = (1 + \varepsilon \operatorname{Tr}(C)) \det \rho$, so keeping the determinant unchanged is equivalent to $\operatorname{Tr}(C) = 0$, that is, C is actually a cocycle valued in $\mathrm{ad}^0 \rho$.

Since we're interested in deformations of global Galois groups, we're also going to be interested in deformations satisfying local conditions. That is, if v is a place of F, there is a homomorphism $G_v \hookrightarrow G$, so by contravariance, we have a restriction map $H^i(G, M) \to H^i(G_v, M)$. This lets us try to understand global cohomology classes in terms of their restrictions to the local Galois groups. For example, we could look at the subgroup of everywhere uramified cohomology classes:

$$\{c \in H^i(G, M) \mid \mathrm{res}_v(c) \text{ is unramified}\}$$

We make the following definition:

Definition 4.1. *Let $\mathcal{L} = (L_v)$ be a collection of subgroups $L_v \subset H^1(G_v, M)$ such that for almost all places v, $L_v = H^1_{nr}(G_v, M)$ (this is called a family of local conditions). The generalized Selmer group is*

$$H^1_{\mathcal{L}}(G_F, M) := \{c \in H^1(G_F, M) \mid \mathrm{res}_v(c) \in L_v \forall v\}$$

We also let $\mathcal{L}^D$ (the dual) denote the family of local conditions (L_v^D), where L_v^D is the annhilator of L_v under the Tate local duality pairing.

Here's an example of a family of local conditions: Fix a finite set $S \supset S_\infty$ of places of a global field F, and let $\rho : G_F \to \mathrm{GL}_n(R)$ be a representation of the absolute Galois group of F. Then we set

- $L_\ell = H^1_{nr}(G_\ell, \mathrm{ad}^0\,\rho)$ if $\ell \notin S$, $\ell \neq p$
- $L_\ell = H^1(G_\ell, \mathrm{ad}^0\,\rho)$ if $\ell \in S$
- L_p the conditions for ordinary deformations

4.2 Global Euler-Poincaré characteristic and Poitou-Tate

The Poitou-Tate nine-term exact sequence is the following: Let F be a number field, and let S be any set of places containing the archimedean places and the places v with $v(\#M) \neq 0$,

$$\begin{aligned} 0 \;&\to\; H^0(F_S, M) \to P^0(F_S, M) \to H^2(F_S, A')^\vee \\ &\to\; H^1(F_S, M) \to P^1(F_S, M) \to H^1(F_S, A')^\vee \\ &\to\; H^2(F_S, M) \to P^2(F_S, M) \to H^0(F_S, A')^\vee \end{aligned}$$

This bears some explanation, since we haven't defined the groups P^i, or the maps in the sequence. Let A be a finite G_F-module. We define

$$P^i(F_S, M) := \prod_{v \in S}{}' H^i(F_v, M)$$

Here the restricted product is taken with respect to the unramified cohomology classes, that is,

$$P^i(F_S, M) = \{(c_v)_{v\in S} \in \prod_{v\in S} H^i(F_v, M) \mid c_v \in H^i_{nr}(F_v, M) \text{ for almost all } v \in S\}$$

Moreover, for archimedean places $v \in S$, we replace H^0 by the modified Tate cohomology group $\hat{H}^0$. In particular,

$$\begin{aligned} P^0(F_S, M) &= \prod_{v\in S\smallsetminus S_\infty} H^0(F_v, M) \times \prod_{v\in S_\infty} \hat{H}^0(F_v, M) \\ P^1(F_S, M) &= \prod_{v\in S}{}' H^1(F_v, M) \\ P^2(F_S, M) &= \bigoplus_{v\in S} H^2(F_v, M) \end{aligned}$$

(by passing to a finite extension where A is unramified).

These groups have topologies: in order, excluding the zero terms, they are finite discrete, compact, compact, discrete, locally compact, compact, discrete, discrete, finite.

Now we want to say what the maps are. The maps $H^i \to P^i$ are evident. For the maps $P^i \to H^{2-i}$, note that local duality gives an isomorphism $P^i \tilde{\to} (P^{2-i})^\vee$ for $0 \le i \le 2$; composing with the (Pontryagin) dual of the homomorphism $H^{2-i} \to P^{2-i}$ gives the desired map. That leaves the maps $(H^2)^\vee \to H^1$ and $(H^1)^\vee \to H^2$. Denoting the maps $H^i \to P^i$ by α_i, there is a non-degenerate pairing $\ker\alpha_1 \times \ker\alpha_2 \to \mathbb{Q}/\mathbb{Z}$, which defines the desired maps.

A theorem due to Poitou and Tate (independently) states that this sequence is exact, and all of the maps are continuous.

Now we would like an analogue of the local Euler-Poincaré characteristic, for global Galois cohomology. We need to assume that S is a finite set, containing S_∞ and the places v with $v(\#M) \neq 0$. First of all, we show that if M is a finite G_S-module, then $H^1(F_S, M)$ is finite. It is also true that $H^i(F_S, A)$ is finite for $i \neq 1$, but this is harder (this is Theorem 8.3.19 in Neukirch)

Proof. We can pass to a finite Galois extension F'/F such that $G_{F',S}$ acts trivially on M. Then $H^1(G_{F',S}, M)$ is finite, because it's equal to $\mathrm{Hom}(F'_S, M)$,

which classifies Galois extensions of F' unramified outside S with Galois group a subgroup of M, and there are finitely many of these (Hermite-Minkowski). Then use the spectral sequence or inflation-restriction to say that $H^1(F_S, M)$ itself is finite. $\square$

Now we define the global Euler-Poincaré characteristic to be

$$\chi(F_S, M) := \frac{\#H^0(F_S, M)\#H^2(F_S, M)}{\#H^1(F_S, M)}$$

We have the formula

$$\chi(F_S, M) = \prod_{v \in S_\infty} \frac{\#H^0(F_v, M)}{\|\#M\|} = \prod_{v \in S_\infty} \frac{\#\hat{H}^0(F_v, M)}{\#H^0(F_v, M')} \tag{4.1}$$

Note that this formula is in terms of the cardinality of the cohomology groups. In this seminar, we will be interested in the case where the cohomology coefficients are vector spaces (either over finite fields or over p-adic fields), so we would like a formula in terms of the dimensions of cohomology groups as vector spaces.

So suppose that M is a finite dimensional vector space over a finite field $k = \mathbb{F}_q$. Then the cohomology groups $H^i(G_S, M)$ are vector spaces over k, so we may take the base q logarithm of 4.1 to get

$$\log_q \chi(F_S, M) = \sum_{v \in S_\infty} \left(h^0(F_v, M) - \log_q \|\#M\|\right) = \sum_{v \in S_\infty} \left(\hat{h}^0(F_v, M) - h^0(F_v, M')\right)$$

5 Product formula

The formula we want to prove is due to Wiles: Let M be a finite G_F-module and let $\mathcal{L}$ be a collection of local conditions. Then

$$\frac{\#H^1_{\mathcal{L}}(F, M)}{H^1_{\mathcal{L}^D}(F, M')} = \frac{\#H^0(F, M)}{\#H^0(F, M')} \cdot \prod_v \frac{\#\mathcal{L}_v}{\#H^0(F_v, M)}$$

where the product runs over all places of F.

We choose a finite set S of places of F as follows: S contains all archimedean places of F, all non-archimedean places whose residue characteristic divides

$\#M$, all places where M is ramified, and all places $\mathfrak{p}$ where $L_\mathfrak{p} \neq \#H^1(G_\mathfrak{p}/I_\mathfrak{p}, M^{I_\mathfrak{p}})$. Let F_S be the maximal extension of F unramified outside S, and let G_S be $\mathrm{Gal}(F_S/F)$.

For any finite discrete G_F-module M, we have an exact sequence

$$0 \to H^1_{\mathcal{L}}(F, M) \to H^1(G_S, M) \to \oplus_{v \in S} H^1(G_v, M)/L_v$$

Taking this exact sequence for M^* and hitting it with $\mathrm{Hom}(-, \mathbb{Q}/\mathbb{Z})$, we get

$$\prod_{v \in S} L_v \to H^1(G_S, M^*)^\vee \to H^1_{\mathcal{L}^D}(F, M^*)^\vee \to 0$$

Here the $\vee$ refers to Pontryagin dual. The identity $(H^1(G_v, M^*)/L_v^D)^\vee = L_v$ follows from local duality: $\mathrm{Hom}(H^1(G_v, M^*)/L_v^D, \mathbb{Q}/\mathbb{Z})$ is the subset of $H^1(G_v, M)$ killing L_v^D under the Tate pairing, which is to say that it is L_v again.

Next we want to merge this exact sequence into the Poitou-Tate exact sequence:

$$\begin{aligned} 0 \quad &\to \quad H^0(G_S, M) \to P^0(G_S, M) \to H^2(G_S, A')^\vee \to \\ &\to \quad H^1_{\mathcal{L}}(F, M) \to \prod_{v \in S} L_v \to H^1(G_S, M^*)^\vee \to H^1_{\mathcal{L}^D}(F, M^*)^\vee \to 0 \end{aligned}$$

If this sequence is exact, we have

$$\frac{\#H^1_{\mathcal{L}}(F, M)}{\#H^1_{\mathcal{L}^D}(F, M')} = \frac{\#H^0(G_S, M)\#H^2(G_S, M')}{\#H^1(G_S, M')\#P^0(G_S, M)} \cdot \prod_{v \in S} \#L_v$$

because $H^2(G_S, M')^\vee$ has the same number of elements as $H^2(G_S, M')$. The formula for $\chi(G_S, M')$ is $\chi(G_S, M') = \prod_{v \in S_\infty} \frac{\hat{h}^0(F_v, M')}{h^0(F_v, M)}$, which yields

$$\begin{aligned} \frac{\#H^1_{\mathcal{L}}(F, M)}{\#H^1_{\mathcal{L}^D}(F, M')} &= \frac{\#H^0(G_S, M)}{\#H^0(G_S, M')} \cdot \frac{1}{\#P^0(G_S, M)} \cdot \prod_{v \in S_\infty} \frac{\#\hat{H}^0(F_v, M')}{\#H^0(F_v, M)} \prod_{v \in S} \#L_v \\ &= \frac{\#H^0(G_S, M)}{\#H^0(G_S, M')} \prod_{v \in S} \frac{\#L_v}{\#H^0(F_v, M)} \text{ by the definition of } P^0 \\ &= \frac{\#H^0(G_S, M)}{\#H^0(G_S, M')} \prod_{v} \frac{\#L_v}{\#H^0(F_v, M)} \end{aligned}$$

The last line follows because outside of S, $L_v = H^1_{nr}(F_v, M)$ and M is unramified, so we can apply the argument that $\#H^1_{nr} = \#H^0$ to say the quotient is 1.

References

[1] Karl Rubin, *Euler Systems*, Princeton University Press, New York: 2000.

[2] J-P. Serre and P. Ion (trans.), *Galois Cohomology*, Springer-Verlag, New York: 1997.

[3] J-P. Serre and M.J. Ion (trans.), *Local Fields*, Springer-Verlag, New York: 1979.

[4] John Tate, "Relations between K_2 and Galois cohomology". *Invent. Math.*. 36 (1976), 257–274.

Lecture 8: Hecke algebras and Galois representations

Burcu Baran
February, 2010

1. $\mathbf{Z}$-Finiteness of Hecke Algebras

Let S_k denote the complex vector space $\mathrm{S}_k(\Gamma_1(N))$ of cusp forms of weight $k \geq 2$ on $\Gamma_1(N)$. Let $\mathbf{T}$ be the $\mathbf{Z}$-subalgebra of $\mathrm{End}_{\mathbf{C}}(\mathrm{S}_k)$ generated by Hecke operators T_p for every prime p and diamond operators $\langle d \rangle$ for every $d \in (\mathbf{Z}/N\mathbf{Z})^\times$. In this section our aim is to prove that $\mathbf{T}$ is a finite free $\mathbf{Z}$-module. As it is clear that $\mathbf{T}$ is torsion-free, it is enough to show that $\mathbf{T}$ is a finitely generated $\mathbf{Z}$-module. We show this in Theorem 1.6.

We begin with some general constructions for any congruence subgroup Γ of $\mathrm{SL}_2(\mathbf{Z})$. Let $\{e, e'\}$ be a $\mathbf{C}$-basis for $\mathbf{C}^2$. The group Γ acts on $\mathbf{C}^2$ via the embedding $\mathrm{SL}_2(\mathbf{Z}) \hookrightarrow \mathrm{SL}_2(\mathbf{C})$ with respect to the basis $\{e, e'\}$: for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$ and $c_1 e + c_2 e' \in \mathbf{C}^2$,

$$\gamma \cdot (c_1 e + c_2 e') = (ac_1 + bc_2)e + (cc_1 + dc_2)e'.$$

This action induces an action on $V_k := \mathrm{Sym}^{k-2}(\mathbf{C}^2)$.

Fix any z_0 in the upper half-plane $\mathfrak{h}$. Let f be any element of the $\mathbf{C}$-vector space $\mathrm{M}_k(\Gamma)$ of modular forms of weight k on Γ. We define the function $I_f : \Gamma \longrightarrow V_k$ by

$$I_f(\gamma) = \int_{z_0}^{\gamma z_0} (ze + e')^{k-2} f(z) dz \tag{1.1}$$

for every $\gamma \in \Gamma$.

Proposition 1.1. *The function I_f in* (1.1) *is a* 1*-cocycle and its class in $\mathrm{H}^1(\Gamma, V_k)$ is independent of z_0.*

Proof. First, we show that I_f is in $\mathrm{Z}^1(\Gamma, V_k)$. Let $\gamma_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and γ_2 be elements of Γ. Since $f|_k \gamma_1 = f$, we have

$$\begin{aligned} \gamma_1 \cdot I_f(\gamma_2) &= \int_{z_0}^{\gamma_2 z_0} ((az+b)e + (cz+d)e')^{k-2} f(z) dz, \\ &= \int_{z_0}^{\gamma_2 z_0} (\gamma_1(z)e + e')^{k-2} f(\gamma_1 z) \frac{dz}{(cz+d)^2}, \\ &= \int_{z_0}^{\gamma_2 z_0} (\gamma_1(z)e + e')^{k-2} f(\gamma_1 z) d(\gamma_1 z), \\ &= \int_{\gamma_1 z_0}^{\gamma_1 \gamma_2 z_0} (ze + e')^{k-2} f(z) dz. \end{aligned} \tag{1.2}$$

It follows that

$$\gamma_1 \cdot I_f(\gamma_2) + I_f(\gamma_1) = \int_{\gamma_1 z_0}^{\gamma_1 \gamma_2 z_0} (ze + e')^{k-2} f(z) dz + \int_{z_0}^{\gamma_1 z_0} (ze + e')^{k-2} f(z) dz = I_f(\gamma_1 \gamma_2),$$

as desired.

Now we show that I_f modulo $\mathrm{B}^1(\Gamma, V_k)$ is independent of z_0. Choose $z_1 \in \mathfrak{h}$. For any $\gamma \in \Gamma$ the difference $\int_{z_0}^{\gamma z_0} (ze + e')^{k-2} f(z) dz - \int_{z_1}^{\gamma z_1} (ze + e')^{k-2} f(z) dz$ is equal to

$$\int_{\gamma z_1}^{\gamma z_0} (ze + e')^{k-2} f(z) dz - \int_{z_1}^{z_0} (ze + e')^{k-2} f(z) dz.$$

The calculations in (1.2) with γz_0 replaced by z_1 show that $\int_{\gamma z_1}^{\gamma z_0} (ze + e')^{k-2} f(z) dz = \gamma \cdot \int_{z_1}^{z_0} (ze + e')^{k-2} f(z) dz$. Hence, we see that the difference is a 1-coboundary.

■

By Proposition 1.1 we can define the $\mathbf{C}$-linear map

$$j : \mathrm{M}_k(\Gamma) \longrightarrow \mathrm{H}^1(\Gamma, V_k) \tag{1.3}$$

by $j(f) = I_f$, where I_f is given in (1.1).

Proposition 1.2. *Choose $z_0 \in \mathfrak{h}$. The restriction*

$$\begin{aligned} j : \mathrm{S}_k(\Gamma) &\longrightarrow \mathrm{H}^1(\Gamma, V_k) \\ f &\mapsto \big(\gamma \mapsto \int_{z_0}^{\gamma z_0} (ze + e')^{k-2} f dz\big), \end{aligned}$$

of (1.3) *is injective.*

Proof. For any $h \in \mathrm{S}_k(\Gamma)$ consider the holomorphic map

$$(ze + e')^{k-2} h(z) : \mathfrak{h} \longrightarrow V_k.$$

Since $\mathfrak{h}$ is simply connected, we can choose a holomorphic function $G_h : \mathfrak{h} \longrightarrow V_k$ so that $dG_h = (ze + e')^{k-2} h(z) dz$. For any $\sigma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z})$ we see that

$$\begin{aligned} d(G_h \sigma) &= G_h'(\sigma(z)) d\sigma(z)\,, \\ &= \big(\big(\frac{az+b}{cz+d}\big)e + e'\big)^{k-2} h(\sigma(z)) \frac{dz}{(cz+d)^2}\,, \\ &= ((az+b)e + (cz+d)e')^{k-2} (h|_k\sigma)(z) dz\,, \end{aligned}$$

where $(h|_k\sigma)(z) = (cz+d)^{-k} h(\sigma(z))$. Therefore, for every $\sigma \in \mathrm{SL}_2(\mathbf{Z})$ we have

$$G_h \sigma = \sigma \cdot G_{h|_k\sigma} + v_\sigma \tag{1.4}$$

for our fixed choice of antiderivative $G_{h|_k\sigma}$ of $(ze + e')^{k-2}(h|_k\sigma)$ and some $v_\sigma \in V_k$.

Let $\mathrm{SL}_2(\mathbf{Z})$ act on the holomorphic maps $G : \mathfrak{h} \longrightarrow V_k$ as follows:

$$(\sigma * G)(z) = \sigma \cdot (G\sigma^{-1}(z)).$$

For each member $\tilde{h}$ of $\mathrm{SL}_2(\mathbf{Z})$-orbit of h (under $\sigma \mapsto h|_k\sigma$) we choose an antiderivative $G_{\tilde{h}}$ as above, so by (1.4) for every $\sigma \in \mathrm{SL}_2(\mathbf{Z})$ we have

$$\sigma * G_h = G_{h|_k\sigma^{-1}} + c_\sigma \tag{1.5}$$

for some $c_\sigma \in V_k$.

Consider $f \in \mathrm{S}_k(\Gamma)$ in the kernel of j; that is, the 1-cocycle

$$\gamma \mapsto \int_{z_0}^{\gamma z_0} (ze + e')^{k-2} f(z) dz = G_f(\gamma z_0) - G_f(z_0)$$

is a 1-coboundary. Then, for every $\gamma \in \Gamma$ we have

$$G_f(\gamma z_0) - G_f(z_0) = \gamma \cdot v - v \tag{1.6}$$

for some $v \in V_k$. Our aim is to show that $f = 0$.

For $\gamma \in \Gamma$ the equation (1.5) becomes

$$\gamma * G_f = G_f + c_\gamma \tag{1.7}$$

for some $c_\gamma \in V_k$. We evaluate this equation at γz_0 and obtain that $c_\gamma = (\gamma * G_f)(z_0) - G_f(\gamma z_0)$. By using equation (1.6) we see that $c_\gamma = \gamma \cdot (G_f(\gamma^{-1} z_0) - v) - (G_f(z_0) - v)$. We may replace G_f with $G_f - (G_f(z_0) - v_\gamma)$, so (1.7) becomes

$$\gamma * G_f = G_f \tag{1.8}$$

for all $\gamma \in \Gamma$.

Recall that for the upper half-plane $\mathfrak{h}$, we topologize $\mathfrak{h}^* = \mathfrak{h} \cup \mathbf{P}^1(\mathbf{Q})$ using $\mathrm{SL}_2(\mathbf{Z})$-translates of bounded vertical strips

$$\{z \in \mathfrak{h} |\, \mathrm{Im}(z) > c,\, a < \mathrm{Re}(z) < b\}$$

for a, $b \in \mathbf{R}$ and $c > 0$. Now we prove the following claim.

Claim 1: As we approach any fixed cusp in $\mathfrak{h}^*$, the function G_f remains bounded in V_k.

Proof of Claim 1: Let $s \in \mathfrak{h}^*$ be any cusp and choose $\sigma \in \mathrm{SL}_2(\mathbf{Z})$ such that $\sigma(s) = \infty$. To prove the claim, it is enough to prove that $\sigma * G_f$ is bounded as we approach ∞ in $\mathfrak{h}$. By (1.5), this is just an antiderivative of $f|_k\sigma^{-1}$. Thus, it suffices to prove that each coefficient function of $(ze+e')^{k-2}(f|_k\sigma^{-1})(z)$ has bounded antiderivative as $\mathrm{Im}(z) \to \infty$ in any bounded vertical strip $\{z \in \mathfrak{h} | \, |\mathrm{Re}(z)| < a\}$ where $a \in \mathbf{R}^+$. Since $f \in \mathrm{S}_k(\Gamma)$, we have $(f|_k\sigma^{-1})(z) \in \mathrm{S}_k(\sigma\Gamma\sigma^{-1})$. Let $\tilde{f}(z) := (f|_k\sigma^{-1})(z)$. Since $\tilde{f}$ is a cusp form for $\sigma\Gamma\sigma^{-1}$, for any $a > 0$ there exists $c \in \mathbf{R}^+$ such that

$$|\bar{f}(z)| \ll e^{-c\mathrm{Im}(z)} \quad \text{as } \mathrm{Im}(z) \to \infty$$

uniformly for $|\mathrm{Re}(z)| < a$. Thus, for any $x \in [-a, a]$ and $y_0 \geq M > 0$ the coefficients of $G_{\tilde{f}}(x+iY) - G_{\tilde{f}}(x+iy_0)$ are linear combinations of terms $\int_{y_0}^{Y} y^r \tilde{f}(x+iy)dy$ with uniformly bounded coefficients. This integral is bounded above by $|P_r(Y)|e^{-cY} + |P_r(y_0)|e^{-cy_0}$, where P_r is a fixed polynomial of degree r, and as $Y \to \infty$ this tends to $|P_r(y_0)|e^{-cy_0}$ uniformly in $|x| \leq a$. This shows that each coefficient function of $(ze+e')^{k-2}(\tilde{f}(z))$ has bounded antiderivative as $\mathrm{Im}(z) \to \infty$ in the mentioned vertical strips. Hence, Claim 1 follows.

Using the $\mathrm{SL}_2(\mathbf{Z})$-invariant bilinear pairing $B : \mathbf{C}^2 \times \mathbf{C}^2 \longrightarrow \mathbf{C}$ defined by the determinant, we obtain the induced bilinear pairing

$$B_k : V_k \times V_k \longrightarrow \mathbf{C},$$

which is also $\mathrm{SL}_2(\mathbf{Z})$-invariant. For $\omega_f = (ze+e')^{k-2} f dz$, consider the 2-form

$$\begin{aligned} B_k(\omega_f, \bar{\omega}_f) &= (k-2)!\,|f|^2 \det(ze+e', \bar{z}e+e')^{k-2}\, dz \wedge d\bar{z}, \\ &= (k-2)!\,(2i)^{k-1}\, y^k\, |f|^2\, \frac{dxdy}{y^2}, \end{aligned} \tag{1.9}$$

where $z = x + iy$. Since f is a cusp form, $B_k(\omega_f, \bar{\omega}_f)$ has finite integral over a fundamental domain F of Γ. Before computing this integral, we compute $B_k(\omega_f, \bar{\omega}_f)$ in another way.

Since $\omega_f = dG_f = g dz$ for $g = (ze+e')^{k-2} f$,

$$B_k(\omega_f, \bar{\omega}_f) = B_k(g, \bar{g})\, dz \wedge d\bar{z}.$$

But g is holomorphic, so $\frac{\partial g}{\partial \bar{z}} = 0$ and hence

$$B_k(g, \bar{g}) = \frac{\partial B_k(G_f, \bar{g})}{\partial z}.$$

Thus, we see that

$$B_k(\omega_f, \bar{\omega}_f) = \frac{\partial B_k(G_f, \bar{g})}{\partial z} dz \wedge d\bar{z} = d(B_k(G_f, \bar{g}) d\bar{z}).$$

By using this equality and Stoke's Theorem we obtain

$$\int_F B_k(\omega_f, \bar{\omega}_f) = \int_{\partial F} B_k(G_f, d\overline{G}_f). \tag{1.10}$$

Now, we want to compute $\int_{\partial F} B_k(G_f, d\overline{G}_f)$. To do this, for each cusp c we choose $\gamma \in \mathrm{SL}_2(\mathbf{Z})$ such that $\gamma(c) = \infty$. We define the "loop" $R_{c,h}$ around c in F to be $\gamma^{-1}(L)$ where L is the horizontal segment joining the two edges at a common "height" h emanating from ∞ in $\gamma(F)$. Define the "closed disc" $D_{c,h} = \gamma^{-1}(U_L)$ where U_L is the closed vertical strip above L including ∞. Then, this integral is equal to

$$\lim_{h \to \infty} \Big(\int_{\partial(F - \cup_c D_{c,h})} B_k(G_f, d\overline{G}_f) + \sum_{c \in \{\text{cusps of } F\}} \int_{R_{c,h}} B_k(G_f, d\overline{G}_f) \Big). \tag{1.11}$$

To calculate the first integral in (1.11) we prove the following claim.

Claim 2: For any $\gamma \in \Gamma$, the pullback $\gamma^*(B_k(G_f, d\overline{G}_f))$ is equal to $B_k(G_f, d\overline{G}_f)$.

Proof of Claim 2: Let $\gamma \in \Gamma$. Since B_k is $\mathrm{SL}_2(\mathbf{Z})$-invariant, we have

$$\gamma^*(B_k(G_f, d\overline{G}_f)) = B_k(G_f\gamma, d(\overline{G_f\gamma})).$$

Since $\gamma \in \Gamma$, by (1.8) we see that $G_f = \gamma^{-1} * G_f$. With this equality we obtain $G_f\gamma = \gamma^{-1} \cdot G_f$. Thus, the above equality gives us

$$\begin{aligned} \gamma^*(B_k(G_f, d\overline{G}_f)) &= B_k(\gamma^{-1} \cdot G_f \,,\, d(\gamma^{-1} \cdot \overline{G}_f)), \\ &= B_k(\gamma^{-1} \cdot G_f \,,\, \gamma^{-1} \cdot d(\overline{G}_f)), \\ &= B_k(G_f, d\overline{G}_f). \end{aligned}$$

The last equality holds because B_k is $\mathrm{SL}_2(\mathbf{Z})$-invariant. Hence, Claim 2 follows.

By Claim 2, the integrals on edges L_1 and L_2 of F such that $L_1 = \gamma L_2$ for some $\gamma \in \Gamma$ cancel. That gives us

$$\int_{\partial(F - \cup_c D_{c,h})} B_k(G_f, d\overline{G}_f) = 0 \tag{1.12}$$

for any h. Now, consider any cusp c of F and the loop $R_{c,h}$ around it. We want to compute $\lim_{h\to\infty} \int_{R_{c,h}} B_k(G_f, d\overline{G}_f)$. Choose $\sigma \in \mathrm{SL}_2(\mathbf{Z})$ such that $\sigma(\infty) = c$. We have

$$\begin{aligned} \int_{R_{c,h}} B_k(G_f, d\overline{G}_f) &= \int_{\sigma^{-1}(R_{c,h})} \sigma^*(B_k(G_f, d\overline{G}_f)), \\ &= \int_{\sigma^{-1}(R_{c,h})} B_k(G_f\sigma, d\overline{G_f\sigma}); \end{aligned} \tag{1.13}$$

the last equality holds because B_k is $\mathrm{SL}_2(\mathbf{Z})$-invariant. The loop $\sigma^{-1}(R_{c,h})$ is a loop $R_{\infty,h}$ around ∞ at height h. By equation (1.4), the function $G_f\sigma$ is just $\sigma \cdot G_{f|_k\sigma}$ up to translation by a constant in V_k. Thus, as B_k is $\mathrm{SL}_2(\mathbf{Z})$-invariant, instead of computing the limit with integral (1.13), we may compute it with $\int_{R_{\infty,h}} B_k(G_{f|_k\sigma}, d\overline{G}_{f|_k\sigma})$ with any choice of antiderivative $G_{f|_k}\sigma$. We do this by calculating the integrals of the $\{e, e'\}$-coefficients of the integrand.

By Claim 1, any antiderivative $G_{f|_k\sigma}$ is bounded in V_k as we approach ∞ in a bounded vertical strip, and $d\overline{G}_{f|_k\sigma}$ has an explicit formula in terms of the cusp form $\bar{f}|_k\sigma$. Thus, for any $a > 0$ there exists $b > 0$ such that

$$|\bar{f}|_k(z)| \ll e^{-b\mathrm{Im}(z)} \quad \text{as } \mathrm{Im}(z) \to \infty$$

uniformly for $|\mathrm{Re}(z)| < a$, so $\lim_{h\to\infty} \int_{R_{\infty,h}} B_k(G_f, d\overline{G_f}) = 0$. As a result, for each cusp c of F and the loop $R_{c,h}$ around it $\lim_{h\to\infty} \int_{R_{c,h}} B_k(G_f, d\overline{G}_f) = 0$. Hence,

$$\lim_{h\to\infty} \sum_{c \in \{\text{cusps of } F\}} \int_{R_{c,h}} B_k(G_f, d\overline{G}_f) = 0. \tag{1.14}$$

By (1.12) and (1.14), we see that the integral (1.10) becomes

$$\int_F B_k(\omega_f, \bar{\omega}_f) = 0.$$

In (1.9), we computed $B_k(\omega_f, \bar{\omega}_f)$ explicitly. Thus, this gives us

$$(k-2)!\,(2i)^{k-1} \int_F y^k\, |f|^2\, \frac{dxdy}{y^2} = 0.$$

The function inside the integral is nonnegative, so $f = 0$, as promised. ∎

From now on, we assume that $\Gamma = \Gamma_1(N)$. By Proposition 1.2, we have injective $\mathbf{C}$-linear map

$$j : \mathrm{S}_k \hookrightarrow \mathrm{H}^1(\Gamma, V_k). \tag{1.15}$$

Now, we want to construct operators acting on $\mathrm{H}^1(\Gamma, V_k)$ compatible via j with the Hecke operators acting on S_k and preserving the $\mathbf{Z}$-structure on $\mathrm{H}^1(\Gamma, V_k)$. To do this we view Hecke operators acting on S_k as double cosets $\Gamma\alpha\Gamma$ where α is an element of

$$\Delta = \{\beta \in \mathrm{M}_2(\mathbf{Z}) \,|\, \det(\beta) > 0,\ \beta \equiv \begin{pmatrix} 1 & * \\ 0 & * \end{pmatrix} \bmod N\}. \tag{1.16}$$

It suffices to construct some T_α acting on $\mathrm{H}^1(\Gamma, V_k)$ for every $\alpha \in \Delta$ such that

(i) the map j in (1.15) carries $[\Gamma\alpha\Gamma]$-action on the left to T_α-action on the right,
(ii) T_α preserves the $\mathbf{Z}$-structure on $\mathrm{H}^1(\Gamma, V_k)$ coming from the one on V_k.

The following three lemmas give such T_α.

Lemma 1.3. *Choose $\alpha \in \Delta$ and coset representatives $\{\alpha_i\}$ for the left multiplication action of Γ in $\Gamma\alpha\Gamma$, so that $\Gamma\alpha\Gamma = \coprod_{i=1}^n \Gamma\alpha_i$. For every i and $\gamma \in \Gamma$, define $j[i]$ uniquely via $\alpha_i\gamma = \gamma_i\alpha_{j[i]}$. There is a well-defined operator*

$$\begin{aligned} T_\alpha : \mathrm{H}^1(\Gamma, V_k) &\longrightarrow \mathrm{H}^1(\Gamma, V_k). \\ c &\longmapsto (\gamma \mapsto \sum_{i=1}^n (\det\alpha)^{k-1}\alpha_i^{-1} \cdot c(\gamma_i)), \end{aligned}$$

which does not depend on the coset representatives.

Let $\Gamma_\alpha := \alpha^{-1}\Gamma\alpha \cap \Gamma$. Using the natural finite-index inclusion $\iota_1 : \Gamma_\alpha \hookrightarrow \Gamma$ and the finite-index inclusion $\iota_2 : \Gamma_\alpha \hookrightarrow \Gamma$ defined by $\iota_2(\beta) = \alpha\beta\alpha^{-1}$, the resulting composite map of the restriction and corestriction maps

$$\mathrm{H}^1(\Gamma, V_k) \xrightarrow[\text{along } \iota_2]{\mathrm{Res}} \mathrm{H}^1(\Gamma_\alpha, V_k) \xrightarrow[\text{along } \iota_1]{\mathrm{Cor}} \mathrm{H}^1(\Gamma, V_k)$$

is the operation T_α.

Proof. We first show that if we use another choice of coset representatives $\{\alpha_i'\}$ for Γ in $\Gamma\alpha\Gamma$, then the operator T_α on 1-cocycles (valued in 1-cochains) changes by 1-coboundaries. Consider

$$\alpha_i' = \tilde{\gamma}_i\alpha_i$$

where $\tilde{\gamma}_i \in \Gamma$ for every i. Since we have $\alpha_i\gamma = \gamma_i\alpha_{j[i]}$ for every i and $\gamma \in \Gamma$, with the new choice of coset representatives we obtain $\tilde{\gamma}_i^{-1}\alpha_i'\gamma = \gamma_i\tilde{\gamma}_{j[i]}^{-1}\alpha_{j[i]}'$. Writing $\gamma_i' := \tilde{\gamma}_i\gamma_i\tilde{\gamma}_{j[i]}^{-1}$, we get

$$\alpha_i'\gamma = \gamma_i'\alpha_{j[i]}'$$

for every i and $\gamma \in \Gamma$. With the new choice of coset representatives $\{\alpha_i'\}$, for $c \in \mathrm{Z}^1(\Gamma, V_k)$ and $\gamma \in \Gamma$ we have the equalities

$$\begin{aligned}
\sum_{i=1}^{n}(\det\alpha)^{k-1}\alpha_i'^{-1}\cdot c(\gamma_i') &= \sum_{i=1}^{n}(\det\alpha)^{k-1}{\alpha_i}^{-1}\tilde{\gamma}_i^{-1}\cdot c(\tilde{\gamma}_i\gamma_i\tilde{\gamma}_{j[i]}^{-1}),\\
&= \sum_{i=1}^{n}(\det\alpha)^{k-1}\alpha_i^{-1}\tilde{\gamma}_i^{-1}\cdot c(\tilde{\gamma}_i)+\sum_{i=1}^{n}(\det\alpha)^{k-1}\alpha_i^{-1}\cdot c(\gamma_i\tilde{\gamma}_{j[i]}^{-1}),\\
&= \sum_{i=1}^{n}(\det\alpha)^{k-1}\alpha_i^{-1}\tilde{\gamma}_i^{-1}\cdot c(\tilde{\gamma}_i)+\sum_{i=1}^{n}(\det\alpha)^{k-1}\alpha_i^{-1}\gamma_i\cdot c(\tilde{\gamma}_{j[i]}^{-1})\\
&\quad+\sum_{i=1}^{n}(\det\alpha)^{k-1}\alpha_i^{-1}\cdot c(\gamma_i),\\
&= -\sum_{i=1}^{n}(\det\alpha)^{k-1}\alpha_i^{-1}\cdot c(\tilde{\gamma}_i^{-1})+\sum_{i=1}^{n}(\det\alpha)^{k-1}\gamma\,\alpha_{j[i]}^{-1}\cdot c(\tilde{\gamma}_{j[i]}^{-1})\\
&\quad+\sum_{i=1}^{n}(\det\alpha)^{k-1}\alpha_i^{-1}\cdot c(\gamma_i),\\
&= \sum_{i=1}^{n}(\det\alpha)^{k-1}\alpha_i^{-1}\cdot c(\gamma_i)+(\gamma\cdot v_0-v_0),
\end{aligned}$$

where $v_0=\sum_{i=1}^{n}(\det\alpha)^{k-1}\alpha_i^{-1}\cdot c(\tilde{\gamma}_i^{-1})$. Hence, we have shown that the operator T_α on 1-cocycles does not depend on the chosen coset representatives if we view its values modulo $\mathrm{B}^1(\Gamma,V_k)$. Now, we want to show that it is a well-defined operator.

We choose coset representatives $\{\alpha_i\}$ for $\Gamma\backslash\Gamma\alpha\Gamma$ so that $\Gamma=\coprod\Gamma_\alpha(\alpha^{-1}\alpha_i)$. We can do this by [1, Lemma 5.1.2]. Since we have $\alpha_i\gamma=\gamma_i\alpha_{j[i]}$ for every $\gamma\in\Gamma$, we see that $(\alpha^{-1}\alpha_i)\gamma=(\alpha^{-1}\gamma_i\alpha)\alpha^{-1}\alpha_{j[i]}$. Since $\alpha^{-1}\alpha_i\in\Gamma$ for every i, we have $(\alpha^{-1}\alpha_i)\gamma(\alpha^{-1}\alpha_{j[i]})^{-1}\in\Gamma$. Thus, it follows from [2, p. 45] that

$$\begin{aligned}
\mathrm{Cor}:\mathrm{H}^1(\Gamma,V_k) &\longrightarrow \mathrm{H}^1(\Gamma_\alpha,V_k),\\
c &\mapsto \big(\gamma\mapsto\sum_{i=1}^{n}(\alpha^{-1}\alpha_i)^{-1}\cdot c((\alpha^{-1}\alpha_i)\gamma(\alpha^{-1}\alpha_{j[i]})^{-1}\\
&\qquad\qquad=\sum_{i=1}^{n}\alpha_i^{-1}\alpha\cdot c(\alpha^{-1}\gamma_i\alpha)\big)
\end{aligned}$$

where $\alpha_i\gamma=\gamma_i\alpha_{j[i]}$. To compute the restriction map along ι_2, observe that the isomorphism

$$\begin{aligned}
V_k &\longrightarrow V_k\\
v &\mapsto \alpha\cdot v
\end{aligned}$$

is equivariant for the Γ_α-action on the left-side and Γ-action on the right-side via the embedding ι_2. Thus, the restriction map is computed as follows

$$\begin{aligned}
\mathrm{Res}:\mathrm{H}^1(\Gamma_\alpha,V_k) &\longrightarrow \mathrm{H}^1(\Gamma,V_k)\\
c &\mapsto \big(\gamma\mapsto\alpha^{-1}\cdot c(\alpha\,\gamma\,\alpha^{-1})\big).
\end{aligned}$$

As a result, we see that the composite map $\mathrm{Cor}\circ\mathrm{Res}$ is the desired map. Hence, T_α is a well-defined action $\mathrm{H}^1(\Gamma,V_k)$.

■

Lemma 1.4. *The T_α-action on $\mathrm{H}^1(\Gamma,V_k)$ is induced by scalar extension of the analogous operation on $\mathrm{H}^1(\Gamma,\mathrm{Sym}^{k-2}(\mathbf{Z}^2))$.*

Proof. Since $k\geq 2$, we have $(\det\alpha)^{k-1}\alpha_i^{-1}=(\det\alpha)^{k-2}((\det\alpha)\,\alpha_i^{-1})$, with $(\det\alpha)\,\alpha_i^{-1}$ a matrix having $\mathbf{Z}$ entries. The result then follows from the cocycle formula for Γ_α.

■

Lemma 1.5. *Consider the action of T_α on $\mathrm{H}^1(\Gamma, V_k)$ that we defined in Lemma 1.3. The injective map j in (1.15) carries the $[\Gamma\alpha\Gamma]$-action on S_k over to the T_α-action on $\mathrm{H}^1(\Gamma, V_k)$ for every α in Δ as in (1.16).*

Proof. Choose $\alpha \in \Delta$ and coset representatives $\{\alpha_i\}$ for $\Gamma\backslash\Gamma\alpha\Gamma$, so $\Gamma\alpha\Gamma = \coprod_{i=1}^n \Gamma\alpha_i$. For $f \in \mathrm{S}_k$ we have $f|_k[\Gamma\alpha\Gamma] = \sum_{i=1}^n f|_k\alpha_i$. Now for each i and $\gamma \in \Gamma$, we compute $I_{f|_k\alpha_i}(\gamma)$ via (1.1):

$$\begin{aligned} I_{f|_k\alpha_i}(\gamma) &= \int_{z_0}^{\gamma z_0} (ze+e')^{k-2}(f|_k\alpha_i)dz, \\ &= \alpha_i^{-1} \cdot \int_{z_0}^{\gamma z_0} \alpha_i \cdot (ze+e')^{k-2}(f|_k\alpha_i)dz, \\ &= \alpha_i^{-1} \cdot (\det \alpha_i)^{k-1} \int_{\alpha_i z_0}^{\alpha_i\gamma z_0} (ze+e')^{k-2} f\, dz. \end{aligned}$$

The last equality follows by the calculations that are similar to the ones that we did in (1.2). Since for $\gamma \in \Gamma$ right multiplication by γ permutes $\Gamma\alpha_i$, for every i and $\gamma \in \Gamma$ there exists a unique $j[i]$ and $\gamma_i \in \Gamma$ such that $\alpha_i\gamma = \gamma_i\alpha_{j[i]}$. By using this equality we compute

$$\begin{aligned} I_{f|_k[\Gamma\alpha\Gamma]}(\gamma) &= (\det\alpha)^{k-1} \sum_{i=1}^n {\alpha_i}^{-1} \cdot \int_{\alpha_i z_0}^{\gamma_i\alpha_{j[i]}z_0} (ze+e')^{k-2} f\, dz, \\ &= (\det\alpha)^{k-1} \sum_{i=1}^n {\alpha_i}^{-1} \cdot \big(\int_{z_0}^{\gamma_i\alpha_{j[i]}z_0} (ze+e') f\, dz - \int_{z_0}^{\alpha_i z_0} (ze+e')^{k-2} f\, dz \big), \\ &= (\det\alpha)^{k-1} \sum_{i=1}^n {\alpha_i}^{-1} \cdot \big(\int_{\gamma_i z_0}^{\gamma_i\alpha_{j[i]}z_0} (ze+e') f\, dz + \int_{z_0}^{\gamma_i z_0} (ze+e') f\, dz \\ &\quad - \int_{z_0}^{\alpha_i z_0} (ze+e') f\, dz \big), \\ &= (\det\alpha)^{k-1} \sum_{i=1}^n {\alpha_i}^{-1} \cdot (\gamma_i \cdot \int_{z_0}^{\alpha_{j[i]}z_0} (ze+e') f\, dz + \int_{z_0}^{\gamma_i z_0} (ze+e') f\, dz \\ &\quad - \int_{z_0}^{\alpha_i z_0} (ze+e') f\, dz \,) \qquad \text{by similar calculations done in (1.2)}, \\ &= (\det\alpha)^{k-1} \big(\sum_{i=1}^n \gamma\, \alpha_{j[i]}^{-1} \cdot \int_{z_0}^{\alpha_{j[i]}z_0} (ze+e') f\, dz + \sum_{i=1}^n \alpha_i^{-1} \cdot \int_{z_0}^{\gamma_i z_0} (ze+e') f\, dz \\ &\quad - \sum_{i=1}^n \alpha_i^{-1} \cdot \int_{z_0}^{\alpha_i z_0} (ze+e') f\, dz \big) \qquad \text{since } \alpha_i^{-1}\gamma_i = \gamma\, \alpha_{j[i]}^{-1}, \\ &= (\det\alpha)^{k-1} \big(\sum_{i=1}^n \alpha_i^{-1} \cdot \int_{z_0}^{\gamma_i z_0} (ze+e') f\, dz + (\gamma \cdot v_1 - v_1) \big), \end{aligned}$$

where $v_1 = \sum_{i=1}^n \alpha_i^{-1} \cdot \int_{z_0}^{\alpha_i z_0} (ze+e') f\, dz$. Therefore, we see that for every $\alpha \in \Delta$ and $f \in \mathrm{S}_k$ we have the quality $j(f|_k[\Gamma\alpha\Gamma]) = T_\alpha(j(f))$ in $\mathrm{H}^1(\Gamma, V_k)$. Hence, the lemma follows. ∎

Theorem 1.6. *Let $\mathbf{T}$ be the $\mathbf{Z}$-subalgebra of $\mathrm{End}_{\mathbf{C}}(\mathrm{S}_k)$ generated by Hecke operators T_p for every prime p and diamond operators $\langle d\rangle$ for every $d \in (\mathbf{Z}/N\mathbf{Z})^\times$. Then $\mathbf{T}$ is finitely generated as a $\mathbf{Z}$-module.*

Proof. By Proposition 1.2, we have $\mathbf{C}$-linear injection

$$j : S_k \longrightarrow \mathrm{H}^1(\Gamma, V_k)$$

for $\Gamma = \Gamma_1(N)$. By Lemma 1.3, for every $\alpha \in \Delta$ (see (1.16)) we have a well-defined action T_α on $\mathrm{H}^1(\Gamma, V_k)$. By Lemma 1.5, the action T_α on $\mathrm{H}^1(\Gamma, V_k)$ is compatible with the action of $[\Gamma\alpha\Gamma]$ on S_k.

Let $\mathbf{T}'$ be the $\mathbf{Z}$-subalgebra of $\mathrm{End}_{\mathbf{C}}(\mathrm{H}^1(\Gamma, V_k))$ generated by the T_α for every $\alpha \in \Delta$. Then, by Lemma 1.4, the $\mathbf{Z}$-algebra $\mathbf{T}'$ is in the image of the $\mathbf{Z}$-subalgebra of $\mathrm{End}_{\mathbf{Z}}(\mathrm{H}^1(\Gamma, \mathrm{Sym}^{k-2}(\mathbf{Z}^2)))$. Since $\mathrm{H}^1(\Gamma, \mathrm{Sym}^{k-2}(\mathbf{Z}^2))$ is a finitely generated $\mathbf{Z}$-module, $\mathbf{T}'$ is also a finitely generated $\mathbf{Z}$-module. By construction, the $\mathbf{T}'$-action on $\mathrm{H}^1(\Gamma, V_k)$ preserves S_k, so we get a restriction map

$$\nu : \mathbf{T}' \longrightarrow \mathrm{End}_{\mathbf{C}}(\mathrm{S}_k)$$

defined by $\nu(T) = T|_{\mathrm{S}_k}$ for every $T \in \mathbf{T}'$. The image of ν in $\mathrm{End}_{\mathbf{C}}(\mathrm{S}_k)$ is $\mathbf{T}$. Therefore, since $\mathbf{T}'$ is finitely generated $\mathbf{Z}$-module, $\mathbf{T}$ is finitely generated $\mathbf{Z}$-module.

■

2. Some Commutative Algebra

In this section we again assume that $\Gamma = \Gamma_1(N)$. Remember that we denote the $\mathbf{C}$-vector space $\mathrm{S}_k(\Gamma_1(N))$ of cusp forms of weight k on Γ by S_k. Let $\mathrm{S}_k(\Gamma, \mathbf{Q})$ be the space of cusp forms with in S_k with Fourier coefficients in $\mathbf{Q}$. By [4, Thm. 3.52], we know that S_k has a $\mathbf{C}$ basis that comes from $\mathrm{S}_k(\Gamma, \mathbf{Q})$ and so we have a surjection

$$\mathrm{S}_k(\Gamma, \mathbf{Q}) \otimes_{\mathbf{Q}} \mathbf{C} \longrightarrow \mathrm{S}_k.$$

Actually, this basis also spans the $\mathbf{Q}$-vector space $\mathrm{S}_k(\Gamma, \mathbf{Q})$ and so this surjection is in fact an isomorphism. This "justifies" the following two definitions.

Definition 2.1. For any field F with characteristic 0,

$$\mathrm{S}_k(\Gamma, F) := \mathrm{S}_k(\Gamma, \mathbf{Q}) \otimes_{\mathbf{Q}} F.$$

Remember that $\mathbf{T}$ is the $\mathbf{Z}$-subalgebra of $\mathrm{End}_{\mathbf{C}}(\mathrm{S}_k)$ generated by Hecke operators T_p for every prime p and diamond operators $\langle d \rangle$ for every $d \in (\mathbf{Z}/N\mathbf{Z})^\times$.

Definition 2.2. For any domain R with characteristic 0, we define

$$\mathbf{T}_R := \mathbf{T} \otimes_{\mathbf{Z}} R$$

acting on $\mathrm{S}_k(\Gamma, \mathrm{Frac}(R))$.

Remark 2.3. By Theorem 1.6 we know that $\mathbf{T}_R$ is a finite free R-module.

Let ℓ be a prime number. Fix an embedding $\overline{\mathbf{Q}} \subset \overline{\mathbf{Q}_\ell}$. Let K be a finite extension of $\mathbf{Q}_\ell$ in $\overline{\mathbf{Q}_\ell}$. Let $\mathcal{O}$ be its ring of integers and λ be its maximal ideal. Consider the finite flat $\mathcal{O}$-algebra $\mathbf{T}_{\mathcal{O}}$.

Proposition 2.4. *The minimal prime ideals of $\mathbf{T}_{\mathcal{O}}$ are those lying over the prime ideal (0) of $\mathcal{O}$.*

Proof. Let P be a minimal prime ideal of $\mathbf{T}_{\mathcal{O}}$. Since $\mathbf{T}_{\mathcal{O}}$ is a flat $\mathcal{O}$-algebra, the going down theorem holds between $T_{\mathcal{O}}$ and $\mathcal{O}$ (see [3, Thm. 9.5]). Therefore, $P \cap \mathcal{O} = (0)$. Now, suppose that P' is a prime ideal of $\mathbf{T}_{\mathcal{O}}$ such that $P' \subset P$ and $P' \cap \mathcal{O} = (0)$. As $\mathbf{T}_{\mathcal{O}}$ is an integral extension of $\mathcal{O}$, there are no strict inclusions between prime ideals lying over (0). Thus, $P' = P$. Hence, the proposition follows. ■

The K-algebra $\mathbf{T}_K$ is Artinian. Hence, it has only a finite number of prime ideals, all of which are maximal. By Proposition 2.4, the natural map

$$\mathbf{T}_{\mathcal{O}} \hookrightarrow \mathbf{T}_{\mathcal{O}} \otimes_{\mathcal{O}} K \cong \mathbf{T}_K$$

induces a bijection

$$\{\text{minimal prime ideals of } \mathbf{T}_{\mathcal{O}}\} \leftrightarrow \{\text{prime ideals of } \mathbf{T}_K\}. \tag{2.1}$$

Moreover, since $\mathcal{O}$ is complete, $T_{\mathcal{O}}$ is λ-adically complete and by [3, Thm. 8.15] there is an isomorphism

$$\mathbf{T}_{\mathcal{O}} \cong \prod_{\mathfrak{m}} \mathbf{T}_{\mathfrak{m}}.$$

The product is taken over the finite set of maximal ideals $\mathfrak{m}$ of $\mathbf{T}_{\mathcal{O}}$ and $\mathbf{T}_{\mathfrak{m}}$ denotes the localization of $\mathbf{T}_{\mathcal{O}}$ at $\mathfrak{m}$. Each $\mathbf{T}_{\mathfrak{m}}$ is a complete local $\mathcal{O}$-algebra which is finite free as an

$\mathcal{O}$-module. With this isomorphism we see that every prime ideal of $\mathbf{T}_{\mathcal{O}}$ is contained in the unique maximal ideal of $\mathbf{T}_{\mathcal{O}}$. Hence, we have a surjection

$$\{\text{minimal prime ideals of } \mathbf{T}_{\mathcal{O}}\} \twoheadrightarrow \{\text{maximal ideals of } \mathbf{T}_{\mathcal{O}}\}. \tag{2.2}$$

Let G_K be the absolute Galois group of K. Suppose $f = \sum a_n q^n$ is a normalized eigenform in $\mathrm{S}_k(\Gamma, \overline{K})$. Then $T \mapsto (T-\text{eigenvalue of } f)$ defines a ring map $\mathbf{T} \longrightarrow \overline{K}$ and so induces a K-algebra homomorphism $\Theta_f : \mathbf{T}_K \longrightarrow \overline{K}$. The image is the finite extension of K generated by the a_n and the kernel is a maximal ideal of $\mathbf{T}_K$ which depends only on the G_K-conjugacy class of f. Thus, we have the map

$$\varphi : \left\{ \begin{array}{c} \text{normalized eigenforms in} \\ \mathrm{S}_k(\Gamma, \overline{K}) \text{ modulo } G_K-\text{conjugacy} \end{array} \right\} \longrightarrow \{\text{maximal ideals of } \mathbf{T}_K\} \tag{2.3}$$

defined by $\varphi(f) = \mathrm{Ker}(\Theta_f)$.

Proposition 2.5. *The map φ in* (2.3) *is a bijection.*

Proof. For any maximal ideal $\mathfrak{m}$ of $\mathbf{T}_K$, all K-algebra embeddings $\mathbf{T}_K/\mathfrak{m} \hookrightarrow \overline{K}$ are obtained from a single one by composing with an element of G_K. Thus, we can make the identification

$$\{\text{maximal ideals of } \mathbf{T}_K\} = \mathrm{Hom}_{K-\mathrm{alg}}(\mathbf{T}_K, \overline{K})/(G_K-\text{action}).$$

Thus, to prove the proposition it is enough to show that the G_K-equivariant map

$$\psi : \{\text{normalized eigenforms in } \mathrm{S}_k(\Gamma, \overline{K})\} \longrightarrow \mathrm{Hom}_{K-\mathrm{alg.}}(\mathbf{T}_K, \overline{K})$$

defined by $\psi(f)(T) = (T-\text{eigenvalue of } f)$ is bijective. To do this, consider the $\overline{K}$-linear map

$$\begin{array}{rcl} \delta : \mathrm{S}_k(\Gamma, \overline{K}) & \longrightarrow & \mathrm{Hom}_{K-\mathrm{vsp}}(\mathbf{T}_K, \overline{K}) \\ f & \mapsto & \big(\alpha_f : T \mapsto a_1(Tf)\big). \end{array} \tag{2.4}$$

If we can show that δ is an isomorphism of $\overline{K}$-vector spaces, then we claim we are done. Because in (2.4) we claim that $f \in \mathrm{S}_k(\Gamma, \overline{K})$ is a normalized eigenform if and only if α_f is a ring homomorphism. To see this, suppose $f \in \mathrm{S}_k(\Gamma, \overline{K})$ is a normalized eigenform, so there exists a K-algebra homomorphism $\Theta_f : \mathbf{T}_K \to \overline{K}$ defined by $Tf = \Theta_f(T)f$ for every $T \in \mathbf{T}_K$. Clearly $\delta(f) = \alpha_f$ where

$$\alpha_f(T) = a_1(Tf) = a_1(\Theta_f(T)f) = \Theta_f(T)a_1(f) = \Theta_f(T)$$

for every $T \in \mathbf{T}_K$. Thus, α_f is a K-algebra homomorphism. Conversely, consider any K-algebra homomorphism $\alpha : \mathbf{T}_K \longrightarrow \overline{K}$, so $\alpha(T) = a_1(Tf)$ for some unique $f \in \mathrm{S}_k(\Gamma, \overline{K})$. Let $\lambda_n = \alpha(T_n)$ for every $T_n \in \mathbf{T}_K$. Then we have

$$a_1(T\,T_n f) = \alpha(T\,T_n) = \alpha(T)\alpha(T_n) = \lambda_n\, a_1(Tf) = a_1(T\,\lambda_n f)$$

for every $T \in \mathbf{T}_K$ and $n \geq 1$. Taking $T = T_m$ for every $m \geq 1$ gives $T_n f = \lambda_n f$ for every $n \geq 1$, proving that f is an eigenform. Moreover, as α is a K-algebra map, $1 = \alpha(\mathrm{id}) = a_1(f)$. Hence, f is a normalized eigenform in $\mathrm{S}_k(\Gamma, \overline{K})$.

Now, we will show that δ is an isomorphism of $\overline{K}$-vector spaces. For injectivity, suppose $\delta(f) = \alpha_f$ is the zero map, so $a_1(Tf) = 0$ for every $T \in \mathbf{T}_K$. In particular, $a_n(f) = a_1(T_n f) = 0$ for every $n \geq 1$, which implies that $f = 0$. To prove surjectivity of δ, it is enough to show that

$$\dim_{\overline{K}} \mathrm{Hom}_{K-\mathrm{vsp}}(\mathbf{T}_K, \overline{K}) \leq \dim_{\overline{K}} \mathrm{S}_k(\Gamma, \overline{K}). \tag{2.5}$$

Since $\mathrm{Hom}_{K-\mathrm{vsp}}(\mathbf{T}_K, \overline{K}) \cong \mathrm{Hom}_{\overline{K}}(\mathbf{T}_{\overline{K}-\mathrm{vsp}}, \overline{K})$, we can work with $\mathrm{Hom}_{\overline{K}}(\mathbf{T}_{\overline{K}-\mathrm{vsp}}, \overline{K})$. Actually, with this identification, studying the map δ is the same as studying the $\overline{K}$-bilinear mapping

$$\begin{array}{rcl} \mathrm{S}_k(\Gamma, \overline{K}) \times \mathbf{T}_K & \longrightarrow & \overline{K} \\ (f \quad , \quad T) & \mapsto & a_1(Tf) \end{array}$$

between finite-dimensional $\overline{K}$-vector spaces. Thus, to prove (2.5), it is enough to show that the map

$$\begin{array}{rcl} \epsilon : \mathbf{T}_{\overline{K}} & \longrightarrow & \mathrm{Hom}_{\overline{K}}(\mathrm{S}_k(\Gamma, \overline{K}), \overline{K}) \\ T & \mapsto & (f \to a_1(Tf)) \end{array}$$

is injective. Suppose $\epsilon(T)$ vanishes for some T. Thus, for every $f \in \mathrm{S}_k(\Gamma, \overline{K})$ and for every integer $n \geq 1$ we have $a_1(T_n T f) = a_1(T T_n f) = 0$. Therefore, $Tf = 0$ for every $f \in \mathrm{S}_k(\Gamma, \overline{K})$. Since $\mathbf{T}_{\overline{K}}$ acts faithfully on $\mathrm{S}_k(\Gamma, \overline{K})$, we get $T = 0$, proving that the map ϵ is injective. Hence, the proposition follows.

■

Combining the bijections (2.1) and (2.3) and the surjection (2.2), we have the following diagram.

$$\begin{array}{ccc} \{\text{minimal prime ideals of } \mathbf{T}_{\mathcal{O}}\} & \twoheadrightarrow & \{\text{maximal ideals of } \mathbf{T}_{\mathcal{O}}\} \\ \updownarrow & & \\ \{\text{prime ideals of } \mathbf{T}_K\} & & \\ \updownarrow & & \\ E = \left\{ \begin{array}{c} \text{normalized eigenforms in} \\ \mathrm{S}_k(\Gamma, \overline{K}) \text{ modulo } G_K-\text{conjugacy} \end{array} \right\} & & \end{array} \tag{2.6}$$

Let $\mathfrak{m}$ be any maximal ideal of $\mathbf{T}_{\mathcal{O}}$, so $\mathfrak{m}$ is the kernel of a map $\Phi : \mathbf{T}_{\mathcal{O}} \longrightarrow \overline{\mathbf{F}}_\ell$. We want to attach a residual representation $\bar{\rho}_{\mathfrak{m}}$ over $\overline{\mathbf{F}}_\ell$ to $\mathfrak{m}$ using the diagram (2.6). Let $\{f_1, ..., f_r\}$ be a set of representatives of all normalized eigenforms in E such that in the diagram (2.6) their corresponding minimal prime ideals $\wp_{f_i}$ in $\mathbf{T}_{\mathcal{O}}$ are inside the maximal ideal $\mathfrak{m}$. For each i, let $\wp'_{f_i}$ be the corresponding prime ideal in $\mathbf{T}_K$, so $\wp'_{f_i} \cap \mathbf{T}_{\mathcal{O}} = \wp_{f_i}$. Thus, for each i, we have a map

$$\begin{array}{rcl} \Theta_{f_i} : \mathbf{T}_{\mathcal{O}} & \longrightarrow & \overline{\mathcal{O}} \\ T_n & \mapsto & a_n(f_i) \end{array}$$

with kernel $\wp_{f_i}$. Since each $\wp_{f_i} \subset \mathfrak{m}$, the map $\Phi : \mathbf{T}_{\mathcal{O}} \longrightarrow \overline{\mathbf{F}}_\ell$ factors through $\mathrm{Im}\,\Theta_{f_i}$ for each i as follows,

$$\begin{array}{ccccc} & & \mathrm{Im}\,\Theta_{f_1} & & \\ & \nearrow & \vdots & \searrow & \\ \mathbf{T}_{\mathcal{O}} & & \longrightarrow & & \overline{\mathbf{F}}_\ell. \\ & \searrow & \vdots & \nearrow & \\ & & \mathrm{Im}\,\Theta_{f_r} & & \end{array}$$

For each i, the quotient $\mathbf{T}_K/\wp'_{f_i}$ is a finite extension K_{f_i} of K. Let $\mathcal{O}_{f_i}$ be its ring of integers and k_{f_i} be its residue field. Each map $\mathrm{Im}\,\Theta_{f_i} \longrightarrow \overline{\mathbf{F}}_\ell$ lifts to $\mathcal{O}_{f_i}$, lifting the embedding of the residue field of $\mathrm{Im}\,\Theta_{f_i}$ to an embedding of k_{f_i} into $\overline{\mathbf{F}}_\ell$. The above commutative diagram tells us that for every integer $n \geq 1$, we have

$$\overline{a_n(f_1)} = \ldots = \overline{a_n(f_r)}$$

in $\overline{\mathbf{F}}_\ell$. Consider the semisimplified residual representation $\bar{\rho}_{f_i}$ associated to each f_i; it is defined over k_{f_i}. For every prime p such that $p \nmid N\ell$ we have

$$\mathrm{tr}(\bar{\rho}_{f_1}(\mathrm{Frob}_p)) = \ldots = \mathrm{tr}(\bar{\rho}_{f_r}(\mathrm{Frob}_p))$$

over $\overline{\mathbf{F}}_\ell$. We obtain a similar result for the determinants of $\bar{\rho}_{f_i}(\mathrm{Frob}_p)$'s when we compare the characters $\bar{\chi}_{f_i}$ associated to f_i's. Therefore, we obtain

$$\bar{\rho}_{f_1} \cong \ldots \cong \bar{\rho}_{f_r}$$

over $\overline{\mathbf{F}}_\ell$. We let $\bar{\rho}_{\mathfrak{m}}$ denote this common residual representation.

3. The Main Theorem

In this section we prove the following theorem.

Theorem 3.1. *Let K be a finite extension of $\mathbf{Q}_\ell$ such that its ring of integers $\mathcal{O}$ is big enough to contain all Hecke eigenvalues at level N. Let λ be its maximal ideal, k its residue field and $\mathfrak{m}$ a maximal ideal of $\mathbf{T}_{\mathcal{O}}$. Consider the associated residual representation*

$$\bar{\rho}_{\mathfrak{m}} : \mathrm{G}_{\mathbf{Q}} \longrightarrow \mathrm{GL}_2(k)$$

over k. Assume $\bar{\rho}_{\mathfrak{m}}$ is absolutely irreducible. Then there exists a unique deformation

$$\rho_{\mathfrak{m}} : \mathrm{G}_{\mathbf{Q}} \longrightarrow \mathrm{GL}_2((\mathbf{T}_{\mathfrak{m}})_{\mathrm{red}})$$

such that

1. *$\rho_{\mathfrak{m}}$ is unramified at every prime p such that $p \nmid N\ell$,*
2. *For every prime p such that $p \nmid N\ell$, the characteristic polynomial of $\rho_{\mathfrak{m}}(\mathrm{Frob}_p)$ is $x^2 - \mathbf{T}_p x + p^{k-1}\langle p \rangle$.*

Before proving this theorem, consider the following theorem which was proved by Akshay in his talk. The corollary of this theorem will be the main ingredient while proving Theorem 3.1.

Theorem 3.2. *Let R be a complete local Noetherian ring and let $\rho : \mathrm{G}_{\mathbf{Q}} \longrightarrow \mathrm{GL}_2(R)$ be a residually absolutely irreducible representation. If S is a complete local Noetherian subring of R which contains all the traces of ρ, then the Galois representation ρ is conjugate to a representation $\mathrm{G}_{\mathbf{Q}} \longrightarrow \mathrm{GL}_2(S)$.*

Corollary 3.3. *Let $\mathcal{O}$ be the ring of integers of a finite extension of $\mathbf{Q}_\ell$, with maximal ideal λ and residue field k. Let Σ be a finite set of places of $\mathbf{Q}$ containing ℓ. Let $\rho : \mathrm{G}_{\mathbf{Q}} \longrightarrow \mathrm{GL}_2(R)$ be the universal deformation unramified outside Σ for an absolutely irreducible representation $\bar{\rho} : \mathrm{G}_{\mathbf{Q}} \longrightarrow \mathrm{GL}_2(k)$ unramified outside Σ, taken on the category of complete local Noetherian $\mathcal{O}$-algebras with residue field k. The traces $\mathrm{tr}(\rho(\mathrm{Frob}_p))$ for all but finitely many primes $p \notin \Sigma$ generate a dense $\mathcal{O}$-subalgebra of R.*

Proof. Let M_R be the maximal ideal of R. By succesive approximation, it is enough to show that such $\mathrm{tr}(\rho(\mathrm{Frob}_p))$ generate $R/(\lambda, M_R^2)$ as k-algebras. Let $R_1 := R/(\lambda, M_R^2)$. The ring R_1 is the universal deformation ring for $\bar{\rho}$ for k-algebras with residue field k such that the square of the maximal ideal is zero. Let S be a k-subalgebra of R_1 generated by $\mathrm{tr}(\rho(\mathrm{Frob}_p))$ for almost all primes $p \notin \Sigma$. Being a subring of R_1, the square of the maximal ideal of S is also zero. If we can show that $R_1 = S$, then we're done.

By Theorem 3.2 we have the following commutative diagram (up to conjugation) which lifts $\bar{\rho}$

$$\begin{array}{ccc} \mathrm{G}_{\mathbf{Q}} & \longrightarrow & \mathrm{GL}_2(S) \\ & \overset{\rho_1}{\searrow} & \downarrow \\ & & \mathrm{GL}_2(R_1) \end{array}$$

Also, since R_1 is the universal deformation ring of $\bar{\rho}$ we have the following commutative diagram (up to conjugation) which lifts $\bar{\rho}$

$$\begin{array}{ccc} \mathrm{G}_{\mathbf{Q}} & \overset{\rho_1}{\longrightarrow} & \mathrm{GL}_2(R_1) \\ & \searrow & \downarrow \\ & & \mathrm{GL}_2(S) \end{array}$$

As a result we have the following composition of maps

$$R_1 \longrightarrow S \hookrightarrow R_1$$

which carries ρ_1 to itself and hence is the identity map. Thus, $S = R_1$. ■

Proof of Theorem 3.1. Let f be a normalized eigenform in $S_k(\Gamma, \overline{K})$ such that the corresponding minimal prime ideal $\mathfrak{p}_f$ in $\mathbf{T}_{\mathcal{O}}$ is contained in $\mathfrak{m}$ (see diagram (2.6)). By Deligne, we have a Galois representation ρ_f over $\mathcal{O}$ associated to f whose residual reduction is $\bar{\rho}_{\mathfrak{m}}$:

$$\begin{array}{ccc} \mathrm{G}_{\mathbf{Q}} & \xrightarrow{\rho_f} & \mathrm{GL}_2(\mathcal{O}) \\ & \searrow^{\bar{\rho}_{\mathfrak{m}}} & \downarrow \\ & & \mathrm{GL}_2(k) \end{array}$$

Let $(R, \rho : \mathrm{G} \longrightarrow \mathrm{GL}_2(R))$ be the universal deformation of $\bar{\rho}_{\mathfrak{m}}$ unramified outside $N\ell$. Then ρ_f corresponds to an $\mathcal{O}$-algebra map $R \longrightarrow \mathcal{O}$, so the diagram

$$\begin{array}{ccc} \mathrm{G}_{\mathbf{Q}} & \xrightarrow{\rho} & \mathrm{GL}_2(R) \\ & \searrow^{\rho_f} & \downarrow \\ & & \mathrm{GL}_2(\mathcal{O}) \end{array}$$

commutes up to conjugation by $1 + \mathrm{M}_2(\lambda)$ in $\mathrm{GL}_2(\mathcal{O})$. By Corollary 3.3, we see that the set of $\mathrm{tr}(\rho(\mathrm{Frob}_q))$ for every prime $q \nmid N\ell$ generates a dense $\mathcal{O}$-subalgebra in R.

Consider the map

$$\begin{array}{rcl} \eta : \ R & \longrightarrow & \displaystyle\prod_{\mathfrak{p}_f \subset \mathfrak{m}} \mathcal{O} \\ \mathrm{tr}(\rho(\mathrm{Frob}_q)) & \mapsto & \displaystyle\prod_{\mathfrak{p}_f} a_q(f) \end{array}$$

where the product is taken over minimal primes $\mathfrak{p}_f$ contained in $\mathfrak{m}$, with f the corresponding normalized eigenform in $\mathrm{S}_k(\Gamma, \overline{K})$. Consider the embedding

$$\begin{array}{rcl} (\mathbf{T}_{\mathfrak{m}})_{\mathrm{red}} & \hookrightarrow & \displaystyle\prod_{\mathfrak{p}_f \subset \mathfrak{m}} \mathbf{T}_{\mathcal{O}}/\mathfrak{p}_f \\ T_q & \mapsto & \displaystyle\prod_{\mathfrak{p}_f} T_q \ (\mathrm{mod}\, \mathfrak{p}_f). \end{array}$$

With the identification

$$\begin{array}{rcl} \displaystyle\prod_{\mathfrak{p}_f \subset \mathfrak{m}} \mathcal{O} & = & \displaystyle\prod_{\mathfrak{p}_f \subset \mathfrak{m}} \mathbf{T}_{\mathcal{O}}/\mathfrak{p}_f \\ \displaystyle\prod_{\mathfrak{p}_f} a_q(f) & \mapsto & \displaystyle\prod T_q \ (\mathrm{mod}\, \mathfrak{p}_f), \end{array}$$

we see that all $\mathrm{tr}(\rho(\mathrm{Frob}_q))$ for $q \nmid N\ell$ land in the closed subalgebra $(\mathbf{T}_{\mathfrak{m}})_{\mathrm{red}}$. Since they generate dense algebra in R, the ring R also lands in there under η, say inducing $h : R \longrightarrow (\mathbf{T}_{\mathfrak{m}})_{\mathrm{red}}$. Thus, we get

$$\rho_{\mathfrak{m}} : \mathrm{G}_{\mathbf{Q}} \xrightarrow{\rho} \mathrm{GL}_2(R) \xrightarrow{h} \mathrm{GL}_2((\mathbf{T}_{\mathfrak{m}})_{\mathrm{red}}).$$

This gives existence and also uniqueness since any other $\rho'_{\mathfrak{m}}$ would give another map $h' : R \longrightarrow (\mathbf{T}_{\mathfrak{m}})_{\mathrm{red}}$ and compatibility with traces of representations then forces $\mathrm{tr}(\rho(\mathrm{Frob}_q)) \mapsto T_q$. Thus, h and h' coincide on a dense set, hence $h = h'$. By checking in each $\mathbf{T}_{\mathcal{O}}/\mathfrak{p}_f = \mathcal{O}$, we see that $\rho_{\mathfrak{m}}(\mathrm{Frob}_q)$ has the expected characteristic polynomial for every $q \nmid N\ell$.

4. Reduced Hecke Algebras

In this section, let K be a finite extension of $\mathbf{Q}_\ell$ and $\mathcal{O}$ its ring of integers. For any ring A, let $\widetilde{\mathbf{T}}_A$ be the A-subalgebra of $\mathbf{T}_A$ generated by the Hecke operators T_p for $p \nmid N\ell$ and diamond operators $\langle d \rangle$ for every $d \in (\mathbf{Z}/N\mathbf{Z})^\times$. Fix a maximal ideal $\mathfrak{m}$ of $\widetilde{\mathbf{T}}_{\mathcal{O}}$. We have a map $\widetilde{\mathbf{T}}_{\mathcal{O}} \longrightarrow \overline{\mathbf{F}}_\ell$ with kernel $\mathfrak{m}$. Since $\mathbf{T}_{\mathcal{O}}$ is an integral extension of $\widetilde{\mathbf{T}}_{\mathcal{O}}$ and $\overline{\mathbf{F}}_\ell$ is algebraically closed, this map can be extended to $\mathbf{T}_{\mathcal{O}}$. Let $\mathfrak{m}'$ be the kernel of this extended map, so it is a maximal ideal of $\mathbf{T}_{\mathcal{O}}$. Consider common (up to isomorphism) residual representation $\bar{\rho}_f$ for all normalized eigenforms f whose corresponding minimal primes $\mathfrak{p}_f$ (see (2.6)) are contained in $\mathfrak{m}'$. Call it $\bar{\rho}_{\mathfrak{m}}$. In this section we prove the following theorem.

Theorem 4.1. *If the Serre conductor $\mathcal{N}(\bar{\rho}_{\mathfrak{m}})$ is equal to N then the $\mathcal{O}$-algebra $(\widetilde{\mathbf{T}}_{\mathcal{O}})_{\mathfrak{m}}$ is reduced.*

Proof. Since the Serre conductor $\mathcal{N}(\bar{\rho}_{\mathfrak{m}})$ is equal to N, the minimal possible level of a normalized eigenform f such that $\bar{\rho}_f \simeq \bar{\rho}_{\mathfrak{m}}$ over $\overline{\mathbf{F}}_\ell$ is N. Thus, such f are newforms. To prove the theorem, we will show that $(\widetilde{\mathbf{T}}_{\mathcal{O}})_{\mathfrak{m}} \otimes_{\mathcal{O}} K$, which contains $(\widetilde{\mathbf{T}}_{\mathcal{O}})_{\mathfrak{m}}$, is reduced. We have the equality

$$(\widetilde{\mathbf{T}}_{\mathcal{O}})_{\mathfrak{m}} \otimes_{\mathcal{O}} K = \prod_{\mathfrak{p}_K} (\widetilde{\mathbf{T}}_K)_{\mathfrak{p}_K}$$

where the product is taken over all prime ideals $\mathfrak{p}_K$ of the Artinian ring $\widetilde{\mathbf{T}}_K$ such that $\mathfrak{p}_K \cap \widetilde{\mathbf{T}}_{\mathcal{O}} \subset \mathfrak{m}$ and $(\tilde{\mathbf{T}}_K)_{\mathfrak{p}_K}$ denotes the localization of $\widetilde{\mathbf{T}}_K$ at $\mathfrak{p}_K$. Thus, each $\mathfrak{p}_K$ in the product corresponds to a newform. To prove the theorem it is therefore enough to show that $(\widetilde{\mathbf{T}}_K)_{\mathfrak{p}}$ is a field when $\mathfrak{p}$ corresponds to a newform.

Assume the prime ideal $\mathfrak{p}$ of $\widetilde{\mathbf{T}}_K$ corresponds to a newform $f \in \mathrm{S}_k(\Gamma, K)$ of level N. We can increase K to a finite extension. Thus, without loss of generality we can assume that K is big enough to contain the Hecke eigenvalues of all normalized eigenforms at level N. Since $\mathrm{S}_k(\Gamma, K)$ is faithful $\widetilde{\mathbf{T}}_K$-module, the localization $(\mathrm{S}_k(\Gamma, K))_{\mathfrak{p}}$ at $\mathfrak{p}$ is faithful $(\widetilde{\mathbf{T}}_K)_{\mathfrak{p}}$-module. If we can prove that $(\mathrm{S}_k(\Gamma, K))_{\mathfrak{p}}$ is one dimensional as a vector space over K then we are done, because this would force $(\widetilde{\mathbf{T}}_K)_{\mathfrak{p}}$ to be equal to K.

We have

$$\mathrm{S}_k(\Gamma, K) = K\, f \oplus \big(\bigoplus_g \mathrm{S}_g(\Gamma, K)\big)$$

where the direct sum is taken over all newforms g of level N_g and $\mathrm{S}_g(\Gamma, K)$ is spanned by $g(vz)$ for the divisors v of N/N_g. By multiplicity one, for every g which is different from f, there exists a prime $q \nmid N\ell$ such that

$$a_q(g(vz)) = a_q(g(z)) \neq a_q(f(z))$$

for every $v|(N/N_g)$. We know that $(T_q - a_q(f)) \in \mathfrak{p}$ and it acts on $g(vz)$ as

$$\begin{aligned} (T_q - a_q(f))g(vz) &= T_q(g(vz)) - a_q(f)g(vz) \\ &= (a_q(g) - a_q(f))g(vz). \end{aligned}$$

By the above argument, $(a_q(g) - a_q(f)) \in K^\times$. But $(\mathbf{T}_K)_{\mathfrak{p}}$ is Artin local, so its maximal ideal is nilpotent. This forces $(\bigoplus_{g \neq f} \mathrm{S}_g(\Gamma, K))_{\mathfrak{p}} = 0$. As a result, $(\mathrm{S}_k(\Gamma, K))_{\mathfrak{p}} = Kf$ and the theorem follows. ∎

References

[1] F. Diamond, J. Shurman, *A first course in modular forms*, Graduate texts in mathematics, **228** (2005), Springer.

[2] S. Lang, *Topics in cohomology of groups*, Lecture notes in mathematics, **1625** (1996), Springer.

[3] H. Matsumura, *Commutative ring theory*, Cambridge studies in advances mathematics , **8** (1986), Cambridge University Press.

[4] G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, Princeton University Press and Iwanami Shoten, (1971), Princeton-Tokyo.

Representations of p-adic groups for the modularity seminar.

Martin H. Weissman

8 January, 2010

Abstract

These are lecture notes, for a "modularity seminar", and I make no claim to originality. I have attempted to give references, but these references do not necessarily reflect the history (I might reference one source for a proof of a theorem, when the theorem was first proven by another). Please send corrections to Marty Weissman at `weissman.marty@gmail.com`.

1 *Notation*

k will always denote a nonarchimedean local field. It will not hurt to assume that $k = \mathbb{Q}_p$.

The valuation on k will be normalized in such a way that $val(k^\times) = \mathbb{Z}$.

$\mathcal{O}$ will always denote the valuation ring of k.

The letter ϖ will always denote a uniformizing element of k, i.e., $val(\varpi) = 1$.

We write $\mathbb{F}_q$ for the residue field of k: $\mathbb{F}_q = \mathcal{O}/\varpi\mathcal{O}$. Here $q = p^f$ for some positive integer f and some prime number p.

We use boldface letters, like $\mathbf{X}$ to denote varieties over k. We use ordinary letters, like X, to denote the k-points of such varieties (with their natural topology).

We often use the language of categories, functors, and natural transformations. In these notes, we typically define functors only half-way: we describe a functor on objects, and leave it to the reader to determine the functor on morphisms when we say something like "For every object X, $F(X)$ is... F extends to a functor from..."'

2 *ℓ-spaces and groups*

Definition 2.1 (Bernstein) *An* ℓ-space[1] *is a locally compact Haus-*

[1] J. Bernstein. *Represenations of p-adic groups*. Harvard University, 1992. Lectures by Joseph Bernstein. Written by Karl E. Rumelhart.

dorff topological space, in which every point has a basis of open compact neighborhoods. Let $\mathfrak{Sp}_\ell$ be the category of ℓ-spaces and continuous maps.

When X is an ℓ-space, the space of "*smooth*" functions on X is defined to be:

$$C^\infty(X) = \{f : X \to \mathbb{C} \colon f \text{ is locally constant}\}.$$

Of course, there is nothing special about $\mathbb{C}$ here – its topology is not being used. Everything we discuss will go through, as long as $\mathbb{C}$ denotes an uncountable, algebraically closed field of characteristic zero.

The subspace $C_c^\infty(X)$ consists of *compactly supported* smooth functions.

Proposition 2.2 *Let X be an ℓ-space, and U an open subset of X with complement $Z = X - U$. Then the linear maps "extension by zero" and "restriction to Z" yield a short exact sequence of complex vector spaces:*

$$0 \to C_c^\infty(U) \to C_c^\infty(X) \to C_c^\infty(Z) \to 0.$$

Example 2.3 *Let $X = k$, where k is a nonarchimedean local field. Let $U = k^\times$ be the open subset of nonzero elements. Then "extension by zero" and "evaluation at zero" yield a short exact sequence of complex vector spaces:*

$$0 \to C_c^\infty(k^\times) \to C_c^\infty(k) \to \mathbb{C} \to 0.$$

Compare and contrast this with the archimedean case – there one should work with Schwarz functions, where one finds that "extension by zero" and "Taylor expansion at 0" yield a short exact sequence of complex vector spaces:[2]

$$0 \to S(\mathbb{R}^\times) \to S(\mathbb{R}) \to \mathbb{C}[[T]] \to 0.$$

[2] Émile Borel. *Sur quelques points de la théorie des fonctions*. Paris., 1894. Original from Columbia University.

The following fact is discussed properly in Chapter 3.1 of Platonov and Rapinchuk[3]:

[3] Vladimir Platonov and Andrei Rapinchuk. *Algebraic groups and number theory*, volume 139 of *Pure and Applied Mathematics*. Academic Press Inc., Boston, MA, 1994. Translated from the 1991 Russian original by Rachel Rowen.

Fact 2.4 *Let $\mathbf{X}$ be an algebraic variety over a nonarchimedean local field k. There is a "natural" topology on $X = \mathbf{X}(k)$ for which X is an ℓ-space. In other words, there is a functor from the category of varieties over k (and regular maps) to the category of ℓ-spaces (and continuous maps), which equals the functor of k-points after forgetting the topology.*

In particular, $\mathbf{GL}_n(k)$ is an ℓ-space, $\mathbf{P}^1(k)$ is an ℓ-space, etc..

In fact, this functor can be uniquely characterized by just a few properties; in unpublished notes[4], Brian Conrad proves:

[4] Brian Conrad. Weil and Grothendieck approaches to adelic points. Unpublished notes, available online.

Theorem 2.5 *Let R be a topological ring. There is a unique functor $\mathbf{X} \mapsto \mathbf{X}(R)$ from the category of affine finite-type R-schemes to the category of topological spaces, such that*

1. *Forgetting the topology yields the functor of R-points.*
2. *The functor is compatible with the formation of fibre products.*

3. *The functor carries closed immersions to topological embeddings.*

4. *The functor applied to* $\mathbf{X} = Spec(R[T])$ *yields the given topology on* $\mathbf{X}(R) = R$.

Furthermore, when R is Hausdorff, closed immersions of schemes yield closed embeddings of topological spaces and when R is locally compact, $\mathbf{X}(R)$ *is locally compact for all* $\mathbf{X}$.

It should be noted that Conrad extends this further, removing the affine hypothesis under the hypothesis that $R^\times$ is open in R, and inversion is continuous on $R^\times$ – these conditions are satisfied when R is a local field.[5]

[5] The situation is more subtle when R is the ring of adeles for a global field; such a ring is locally compact and Hausdorff, but $R^\times$ is no longer open in R.

Definition 2.6 *An* ℓ-group *is a group in the category of* ℓ*-spaces. In other words, an* ℓ*-group is a group G, endowed with a topology for which G is an* ℓ*-space and the unit, inverse, and composition maps:*

$$pt \to G, \quad G \to G, \quad G \times G \to G$$

are continuous.

Proposition 2.7 *Let G be a topological group. Then G is an* ℓ*-group*[6] *if and only if the identity element has a basis of neighborhoods consisting of open compact subgroups of G.*

[6] This is given by some authors as the definition of an ℓ-group. I find it more natural to think about groups in a category and prove the equivalence.

Proof: If G has a neighborhood basis around the identity consisting of open compact subgroups, then translation of these open compact subgroups gives a neighborhood basis around any point in G. It follows quickly that G is an ℓ-space.

Conversely, if G is an ℓ-space and a topological group, then there is a neighborhood basis of the identity consisting of open compact *subsets* of G. Let V be such a compact open subset containing the identity of G. Define

$$K_V = \{x \in G \colon xV \subset V \text{ and } x^{-1}V \subset V\}.$$

Then K_V is a subgroup of G, and a subset of V. It is the intersection of compact sets, hence compact. The proof that K_V is open is a bit tricky, and we refer to the notes of Paul Garrett [7].

[7] P. Garrett. Smooth representations of totally disconnected groups. Introductory notes, available online. Updated July 8, 2005.

Q.E.D

Corollary 2.8 *If* $\mathbf{G}$ *is an algebraic group over a nonarchimedean local field k, then* $G = \mathbf{G}(k)$ *is naturally a* ℓ*-group.*

Here are a few examples of ℓ-groups arising as $\mathbf{G}(k)$, and open compact neighborhoods of the identity.

Example 2.9 *Let $\mathbf{G}_a$ denote the additive group over k. Thus $G_a = (k, +)$ is the additive group of the field k. Let $val : k^\times \to \mathbb{Z}$ denote the valuation on k, normalized to be surjective. For $m \in \mathbb{Z}$, define a compact open subgroup of G_a:*

$$K_m = \{x \in k \colon val(x) \geq m\}.$$

Then

$$G_a = \bigcup_{m \in \mathbb{Z}} K_m, \quad \{0\} = \bigcap_{m \in \mathbb{Z}} K_m.$$

Note above that the additive group G_a is the union of its compact open subgroups. This is not typical, for ℓ-groups. But it does hold for groups $G = \mathbf{G}(k)$, whenever $\mathbf{G}$ is a unipotent group over a p-adic field k. This plays a very important role for harmonic analysis on unipotent p-adic groups.

Example 2.10 *Let $\mathbf{G}_m$ denote the multiplicative group over k. Thus $G_m = k^\times$ is the multiplicative group of the field k. A choice of uniformizing element $\varpi \in k^\times$ determines a decomposition of topological groups:*

$$k^\times \cong \mathcal{O}^\times \times \mathbb{Z}.$$

The compact open subgroups

$$U_m = \{x \in k^\times \colon val(x - 1) \geq m\},$$

for $m \geq 1$, form a neighborhood basis at the identity in $k^\times$.

Of course, $\mathbf{G}_m = \mathbf{GL}_1$, and the above example generalizes to $\mathbf{GL}_n$ without much difficulty.

Example 2.11 *Let $\mathbf{GL}_n$ be the algebraic group of n by n invertible matrices. Let ϖ be a uniformizing element of k. A neighborhood basis of the identity in $GL_n = \mathbf{GL}_n(k)$, consisting of compact open subgroups, is given by:*

$$K_m = \{g \in GL_n(\mathcal{O}_k) \colon g \equiv 1, \text{ modulo } \varpi^n \mathcal{O}_k\}.$$

3 *Representations*

Smooth representations

Let G be an ℓ-group. Nothing will really be lost if one takes $G = \mathbf{GL}_n(\mathbb{Q}_p)$ in what follows.

Definition 3.1 *A* representation *of G is a pair (π, V), where V is a complex vector space (often infinite-dimensional!) and $\pi : G \to Aut_{\mathbb{C}}(V)$ is an action of G on V by $\mathbb{C}$-linear automorphisms. Let $\mathfrak{Rep}_G$ be the category of representations of G and G-intertwining $\mathbb{C}$-linear maps.*

Let $Op(G)$ be the set of open subgroups of G – recall that $Op(G)$ is a basis of neighborhoods of the identity in G. For any subgroup $H \subset G$, and any representation (π, V) of G, we write V^H for the H-invariant subspace of V. We write V_H for the H co-invariant quotient[8] of V, i.e.,

$$V_H = {}^V/_{[H-1]V}, \quad [H-1]V = Span_{\mathbb{C}}\{\pi(h)v - v\}_{v\in V, h\in H}.$$

[8] Let W be a vector space upon which H acts trivially. Then every H-intertwining map from W to V factors uniquely through V^H. Every H-intertwining map from V to W factors uniquely through V_H.

Definition 3.2 *When (π, V) is a representation of G, the subspace V^∞ of smooth vectors is defined by:*

$$V^\infty = \bigcup_{H\in Op(G)} V^H.$$

A representation (π, V) of G is called smooth *if $V = V^\infty$. Let $\mathfrak{Rep}_G^\infty$ denote the category[9] of smooth representations of G and G-intertwining $\mathbb{C}$-linear maps.*

[9] The category $\mathfrak{Rep}_G^\infty$ is an abelian category with enough injectives and arbitrary direct sums.

Proposition 3.3 *If (π, V) is a representation of G, then (π, V^∞) is a subrepresentation[10] of (π, V), and (π, V^∞) is smooth. This defines a functor from $\mathfrak{Rep}_G$ to $\mathfrak{Rep}_G^\infty$. If (σ, W) is any smooth representation of G, and $\phi : W \to V$ is a morphism in $\mathfrak{Rep}_G$, then ϕ factors uniquely through the inclusion $V^\infty \hookrightarrow V$.*

[10] A subrepresentation of (π, V) is just a G-stable subspace.

Proof: The proof is not difficult, and is left to the reader.

Q.E.D

The category $\mathfrak{Rep}_G^\infty$ is usually not semisimple. However, for compact groups the category is semisimple and we discuss this a bit further.

Let K be a compact ℓ-group. Let $\hat{K}$ be a set of representatives for the isomorphism classes of irreducible smooth representations (abbreviated *irrep* hereafter) of K – in other words, if τ is an irrep of K then there exists a unique $\rho \in \hat{K}$ such that $\tau \cong \rho$.

Lemma 3.4 *Every irrep τ of K is finite-dimensional and factors through a finite quotient of K.*

Proof: Let (τ, W) be an irrep of K, and let w be a nonzero vector in W. Let $H \subset K$ be an open subgroup such that $w \in W^H$. By compactness of K, we find that $\#(K/H) < \infty$. Choosing representatives $k_1, \ldots, k_d$ for K/H, we find that

$$\mathrm{Span}_{\mathbb{C}}\{\tau(k)w\}_{k\in K} = \mathrm{Span}_{\mathbb{C}}\{\tau(k_i)w\}_{1\le i\le d}.$$

By irreducibility, the left side is all of W. The right side is finite-dimensional, and so $dim(W) \le d = \#(K/H)$.

Now, each vector $\tau(k_i)w$ is fixed by the open subgroup $k_i H k_i^{-1}$. Hence we find that

$$\tau(k_i)w \in W^N, \text{ where } N = \bigcap_{i=1}^{d} k_i H k_i^{-1}.$$

Observe that N is an open normal subgroup of K, so K/N is a finite quotient of K, and τ factors through this quotient.

Q.E.D

Definition 3.5 *Let (π, V) be a smooth representation of K, and $(\tau, W) \in \hat{K}$. The τ-isotypic subrepresentation of V is the image V_τ of the natural injective K-intertwining operator:*

$$W \otimes_{\mathbb{C}} Hom_K(W, V) \to V.$$

The τ-isotypic subrepresentations of a smooth representation (π, V) of K are certainly semisimple – they are isomorphic to a direct sum of copies of τ.

Theorem 3.6 *Let (π, V) be a smooth representation of K. Then the inclusions of isotypic subrepresentations yield an isomorphism*

$$\bigoplus_{\tau \in \hat{K}} V_\tau \cong V.$$

Proof: Schur's orthogonality (for finite groups) implies that the distinct isotypic subrepresentations of V have zero intersection. Thus it remains to prove that every vector $v \in V$ can be expressed as a finite sum of vectors in isotypic subrepresentations.

But if $v \in V$, then $v \in V^H$ for some open subgroup $H \subset K$. With the techniques of the previous lemma, we find that $v \in V^N$ for some open normal subgroup $N \subset H \subset K$. Let $W \subset V$ be the smallest subrepresentation of K containing v. We find that W is finite-dimensional, and the representation of K on W factors through the quotient K/N.

From the complete decomposability of representations of finite groups, we find that W decomposes into a finite number of K/N-isotypic components. Pulling back, we find that W decomposes into a finite number of K-isotypic components. In particular, v can be expressed as a finite sum of vectors from isotypic subrepresentations of V.

Q.E.D

It is important to contrast the case of compact ℓ-groups (which are really no more difficult than finite groups) with noncompact ℓ-groups. The simplest example of a noncompact ℓ-group is $\mathbb{Z}$ –

every representation of $\mathbb{Z}$ is smooth. The category of representations of $\mathbb{Z}$ is isomorphic to the category of $\mathbb{C}[T^{\pm 1}]$-modules, transferring the action of $n \in \mathbb{Z}$ to the action of $T^n \in \mathbb{C}[T^{\pm 1}]$.

There are plenty of examples of non-semisimple representations of $\mathbb{Z}$; one may take $(\pi, \mathbb{C}^2)$ for example, where

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

There is a short exact sequence of $\mathbb{Z}$-representations:

$$0 \to \mathbb{C} \to (\pi, \mathbb{C}^2) \to \mathbb{C} \to 0,$$

where we write $\mathbb{C}$ here for the trivial representation. This is essentially the best we can do for "decomposing" the representation π into irreducibles.

One might also consider an infinite-dimensional representation, like the space $V = C_c^\infty(\mathbb{Z})$ of compactly (finitely) supported $\mathbb{C}$-valued functions on $\mathbb{Z}$, on which $\mathbb{Z}$ acts by translation π:

$$[\pi(n)f](x) = f(x+n).$$

Then (π, V) has *no irreducible subrepresentation*, though it has infinitely many irreducible quotients. Indeed, summation yields a trivial irreducible quotient

$$\Sigma : V \to \mathbb{C}, \quad \Sigma(f) = \sum_{n \in \mathbb{Z}} f(n).$$

In fact, one can show

Theorem 3.7 *Let (π, V) be a representation of $\mathbb{Z}$. If V is finitely-generated as a $\mathbb{C}[T^{\pm 1}]$-module, then there exists an irreducible quotient of V.*

Proof: Consider V as a $\mathbb{C}[T^{\pm 1}]$-module. Every irreducible representation of $\mathbb{Z}$ is a character (one-dimensional) $\chi_z : \mathbb{Z} \to \mathbb{C}^\times$ (this will follow from Schur's lemma, proven a bit later), for some $z \in \mathbb{C}^\times$, where we define

$$\chi_z(n) = z^n.$$

If $Hom(V, \chi_z) = 0$, then we find that $V/m_z V = 0$ for every maximal ideal $m_z = \langle T - z \rangle$ of $\mathbb{C}[T^{\pm 1}]$. From Nakayama's lemma, it follows that $V = 0$.

Q.E.D

Thus the moral is: smooth representations of noncompact groups often do not have irreducible subrepresentations; but usually (assuming a finite-type hypothesis) have irreducible quotients. Another example of this phenomenon is given by the following

Example 3.8 *Let $V = C_c^\infty(k)$ be the space of smooth (i.e., locally constant) compactly supported functions on k, viewed as a representation of $k = \mathbf{G}_a(k)$ by translation:*

$$[\pi(g)f](x) = f(x+g), \text{ for all } g \in k, x \in k, f \in V.$$

Then V has no irreducible subrepresentation. Indeed, we will see that all irreducible subrepresentations are characters – but if translation of a function acts as a character, the function cannot be compactly supported. However, every irreducible smooth representation of k occurs as a quotient; if $(\psi, \mathbb{C})$ is a smooth character of k then the following gives a nontrivial k-intertwining map from (π, V) to $(\psi, \mathbb{C})$:

$$f \mapsto \int_k f(x)\overline{\psi(x)}dx,$$

where we fix the Haar measure for which $\mathcal{O}$ has measure 1.

Contragredience, admissibility

When (π, V) is a smooth representation of G, the linear dual space $V' = Hom_{\mathbb{C}}(V, \mathbb{C})$ is a representation of G via:

$$[\pi'(g)\lambda)](v) = \lambda(\pi(g^{-1})v) \text{ for all } \lambda \in V', v \in V.$$

But this representation is rarely smooth:

Definition 3.9 *If (π, V) is a smooth representation of G, define $\tilde{V} = (V')^\infty$ – the space of smooth vectors in the linear dual of V. Let $\tilde{\pi}$ denote the resulting representation of G on $\tilde{V}$. The representation $(\tilde{\pi}, \tilde{V})$ is called the* contragredient *representation of (π, V). The contragredient defines a contravariant functor from $\mathfrak{Rep}_G^\infty$ to itself.*

It is very important to note that the contragredient does *not* define a duality – there is a natural transfomation of functors from the identity functor to the double-contragredient, but this is not a natural isomorphism. The contragredient functor does define a duality for *admissible representations*:

Definition 3.10 *A representation (π, V) of G is called* admissible *if it is smooth and for all $H \in Op(G)$, $dim(V^H) < \infty$.*

We may characterize admissible representations also as follows:

Proposition 3.11 *Let (π, V) be a smooth representation of G. Let K be a compact open subgroup of G. Then (π, V) is admissible if and only if for every $\tau \in \hat{K}$, the (K, τ)-isotypic component V_τ is finite-dimensional.*

PROOF: Suppose first that every (K, τ)-isotypic component of V is finite-dimensional. Let H be an open subgroup of G. Let $H' = H \cap K$; then K/H' is finite. Define

$$J = \bigcap_{k \in K/H'} kH'k^{-1}.$$

Then we find that J is a normal subgroup of K, J is open compact, and $J \subset H$.

It follows that $V^H \subset V^J$ and:

$$V^J = \bigoplus_{\tau} (V_\tau)^J.$$

But there are only finitely many isomorphism classes of irreducible smooth representations of K for which $V^J \neq 0$, since there are only finitely many isomorphism classes of irreducible representations of the quotient group K/J. Hence V^J is a finite direct sum, of finite-dimensional spaces. Hence V^J is finite-dimensional, and so V^H is finite-dimensional. Hence V is admissible.

The converse is similar, and left to the reader.

Q.E.D

Proposition 3.12 *Let (π, V) be an smooth representation of G. Then (π, V) is admissible if and only if the natural homomorphism $V \to \widetilde{\widetilde{V}}$ is an isomorphism.*

PROOF: If (π, V) is admissible, one may choose an open compact subgroup $K \subset G$, and decompose V into its isotypic components:

$$V = \bigoplus_{\tau \in \hat{K}} V_\tau.$$

The linear dual of V is then a direct product of finite-dimensional spaces:

$$V' = \prod_{\tau \in \hat{K}} Hom(V_\tau, \mathbb{C}).$$

One may check that the smooth vectors in V' are now:

$$\tilde{V} = \bigoplus_{\tau \in \hat{K}} V'_\tau.$$

It follows that $\tilde{V}$ is admissible.

The details and other converse are left to the reader.

Q.E.D

The following theorem is much deeper.

Theorem 3.13 (Jacquet) [11] *If* $\mathbf{G}$ *is a connected reductive algebraic group over a local nonarchimedean field, and* (π, V) *is an irreducible smooth representation of* $G = \mathbf{G}(k)$*, then* (π, V) *is admissible.*

[11] Hervé Jacquet. Sur les représentations des groupes réductifs *p*-adiques. *C. R. Acad. Sci. Paris Sér. A-B*, 280:Aii, A1271–A1272, 1975.

Corollary 3.14 *If* $\mathbf{G}$ *is a connected reductive algebraic group over a local nonarchimedean field, and* (π, V) *is an irreducible smooth representation of* G *(or a representation of finite length), then* (π, V) *is admissible and* V *is isomorphic to its double contragredient.*

Schur's lemma

Theorem 3.15 (Jacquet) [12] *Suppose that* G *has a countable basis for its topology. Let* (π, V) *be an irreducible smooth representation of* G*. Then the dimension of* V *is countable and* $End_G(V) = \mathbb{C}$*.*

[12] Hervé Jacquet. Sur les représentations des groupes réductifs *p*-adiques. *C. R. Acad. Sci. Paris Sér. A-B*, 280:Aii, A1271–A1272, 1975.

[13] S. DeBacker. *Some notes on the representation theory of reductive p-adic groups*. Available online.

Proof: (We have followed DeBacker's notes[13]) Let $0 \neq v \in V$, and let K be a compact open subgroup of G for which $v \in V^K$. Then G/K is a countable set (since G has a countable basis for its topology) and we may choose representatives $g_1, g_2, \ldots$ for this countable set of cosets. We find that

$$\mathrm{Span}_{\mathbb{C}}\{\pi(g)v\}_{g\in G} = \mathrm{Span}_{\mathbb{C}}\{\pi(g_i)v\}_{i=1,2,\ldots}.$$

The left side is a nonzero subrepresentation of V, hence equals V by irreducibility. The right side is a countable-dimensional vector space, and the first assertion is proven.

For the second assertion, consider any $e \in End_G(V)$, and a nonzero vector $v \in V$ again. The operator e is uniquely determined by $e(v)$, since $e(\pi(g)v) = \pi(g)e(v)$, and the vectors $\pi(g)v$ span V as a complex vector space.

It follows that the map $e \mapsto e(v)$ is an injective $\mathbb{C}$-linear map from $End_G(V)$ to V. Hence $End_G(V)$ has countable dimension. But since V is an irreducible representation of G, we know that $End_G(V)$ is a skew-field. Consider the (commutative) subfield:

$$\mathbb{C}(e) \subset End_G(V).$$

If $\mathbb{C} \neq \mathbb{C}(e)$ – i.e., if e is not a scalar endomorphism of V – then e must be transcendental over $\mathbb{C}$. But note that $\mathbb{C}(e)$ is uncountable-dimensional as a $\mathbb{C}$-vector space since the set

$$\{(e-c)^{-1} : c \in \mathbb{C}\}$$

is uncountable and $\mathbb{C}$-linearly independent. This is a contradiction.

Hence $\mathbb{C} = \mathbb{C}(e)$ – every element of $End_G(V)$ is a scalar endomorphism.

Q.E.D

This adaptation of Schur's lemma has the usual consequences:

Corollary 3.16 *If G is an abelian ℓ-group with countable basis for its topology, then every irreducible representation of G is one-dimensional.*[14]

[14] We call a one-dimensional smooth representation a *character*.

Corollary 3.17 *Let G be an ℓ-group with countable basis for its topology. Let (π, V) be an irreducible smooth representation of G. Let Z be the center of G. Then there exists a smooth character $\chi : Z \to \mathbb{C}^\times$ such that*

$$\pi(z)v = \chi(z) \cdot v \text{ for all } z \in Z, v \in V.$$

When G is an ℓ-group with countable basis for its topology, and center Z, it is often convenient to consider not the category $\mathfrak{Rep}_G^\infty$, but rather the full subcategory consisting of representations with a given central character. If $\chi : Z \to \mathbb{C}^\times$ is a character of Z, and (π, V) is any smooth representation of G, we say that (π, V) has central character χ if $\pi(z)v = \chi(z) \cdot v$ for all $z \in Z$. Of course, not all smooth representations of G have a central character (though irreps do). We define $\mathfrak{Rep}_{G,\chi}^\infty$ to be the full subcategory of $\mathfrak{Rep}_G^\infty$, whose objects are those smooth representations with central character χ.

Corollary 3.18 *If (π, V) and (σ, W) are two irreducible smooth representations of G – an ℓ-group with countable basis for its topology – then $Hom_G(V, W)$ is either zero or one-dimensional.*

Induction, Compact Induction

Our treatment of smooth induction follows Bernstein [15], to some extent. Let H be a closed subgroup of an ℓ-group G. Let (π, V) be a smooth representation of G, and let (σ, W) be a smooth representation of H. Restriction of representations is quite simple:

[15] J. Bernstein. *Represenations of p-adic groups*. Harvard University, 1992. Lectures by Joseph Bernstein. Written by Karl E. Rumelhart.

Definition 3.19 *Define*[16] *$Res_H^G \pi$ to be the restriction of π to H. This extends to a functor, Res_H^G from $\mathfrak{Rep}_G^\infty$ to $\mathfrak{Rep}_H^\infty$.*

[16] We always put the smaller group below, and larger group above, in our notation for induction and restriction.

Induction of representations, as usual, is not as simple.

Definition 3.20 *Define $\mathbb{C}[[H\backslash_\sigma G, W]]$ to be the vector space of functions $f : G \to W$ such that:*

$$f(hx) = \sigma(h)(f(x)), \text{ for all } x \in G, h \in H.$$

This is a representation of G by right translation:

$$[gf](x) = f(xg) \text{ for all } x, g \in G.$$

Define $Ind_H^G W$ to be the subspace $\mathbb{C}[[H\backslash_\sigma G, W]]^\infty$ of smooth vectors for this action. This extends to a functor, Ind_H^G from $\mathfrak{Rep}_H^\infty$ to $\mathfrak{Rep}_G^\infty$.

More concretely, an element of $Ind_H^G W$ is a function $f : G \to W$ which satisfies the following conditions:

1. $f(hx) = \sigma(h)(f(x))$ for all $x \in G, h \in H$.
2. There exists an open subgroup $K \subset G$ such that $f(xk) = f(x)$ for all $x \in G$. In other words, f is *uniformly*[17] *locally constant*.

[17] The uniformity is that K can be chosen independently of x.

There is an important *subfunctor* of Ind_H^G, called *compact induction*:

Definition 3.21 *Define $ind_H^G W$ to be the subspace of $Ind_H^G W$, consisting of those functions $f \in Ind_H^G W$ satisfying the additional condition:*

> *There exists a compact subset $X \subset G$ such that $f(g) = 0$ unless $g \in H \cdot X$. In other words, f is compactly supported, modulo H.*

Then $ind_H^G W$ is a G-subrepresentation of $Ind_H^G W$; it yields a subfunctor $ind_H^G \subset Ind_H^G$.

Compact induction is simpler in many ways; for example, the condition of uniform local constancy simplifies to the condition of local constancy. Of course, if $H\backslash G$ is a compact space, then the functors ind_H^G and Ind_H^G coincide. Less trivially,

Proposition 3.22 *If (σ, W) is an admissible representation of H, and $H\backslash G$ is compact, then $Ind_H^G W$ is an admissible representation of G.*

Proof: We leave the proof as an exercise. This can be found in Proposition 9 of Bernstein's notes[18] as well.

[18] J. Bernstein. *Represenations of p-adic groups*. Harvard University, 1992. Lectures by Joseph Bernstein. Written by Karl E. Rumelhart.

Q.E.D

Frobenius reciprocity can now be formulated in the smooth setting:

Theorem 3.23 *Let (π, V) be a smooth representation of G, and (σ, W) a smooth representation of H, a closed subgroup of G. Then there is a natural isomorphism:*

$$Hom_G(V, Ind_H^G W) \cong Hom_H(Res_H^G V, W).$$

This identifies Ind_H^G as a functor which is right adjoint to the functor Res_H^G. Both functors are exact.

Most typically, the functor ind_H^G of compact induction is used when H is a closed and open (clopen) subgroup of G; in this case, $H\backslash G$ is a discrete space. It follows that

Lemma 3.24 *Let H be a clopen subgroup of G. Let (σ, W) be a smooth representation of H. Then there is a natural isomorphism of representations of G:*

$$ind_H^G W \cong W \otimes_{\mathbb{C}[H]} \mathbb{C}[G].$$

From the adjointness of ring extension and pullback, we find that

Theorem 3.25 (p. 125 of Cartier) [19] *Let H be a clopen subgroup of G. Let (π, V) be a smooth representation of G, and (σ, W) a smooth representation of H, a closed subgroup of G. Then there is a natural isomorphism:*

$$Hom_G(ind_H^G W, V) \cong Hom_H(W, Res_H^G V,).$$

This identifies ind_H^G as a functor which is left adjoint to the functor Res_H^G. Both functors are exact.

[19] P. Cartier. Representations of p-adic groups: a survey. In *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 1*, Proc. Sympos. Pure Math., XXXIII, pages 111–155. Amer. Math. Soc., Providence, R.I., 1979.

Pullback, corestriction

Suppose now that $B = T \ltimes U$, where T, U are closed subgroups of an ℓ-group B. Let $p : B \to T$ be the projection map. There is a functor given by *pullback*:

Definition 3.26 *Let (η, Y) be a smooth representation of T. Define $p^*\eta : B \to Aut_{\mathbb{C}}(V)$ by*

$$p^*\eta(b) = \eta(p(b)).$$

Then $(p^*\eta, Y)$ is a smooth representation of B, and p^* extends to a functor from $\mathfrak{Rep}_T^\infty$ to $\mathfrak{Rep}_B^\infty$.

Of course, one may introduce the general pullback of smooth representations, including restriction to a subgroup as well as the above pullback as special cases. The *pushforward* functor is defined by coinvariants:

Definition 3.27 *Let (σ, W) be a smooth representation of B. Define $p_*W = W_U = W/[U-1]W$ to be the space of U-coinvariants of W. Then, since T normalizes U, it follows that $\sigma(T)$ stabilizes $[U-1]W$ and hence the action σ of T on W descends to an action $p_*\sigma$ of T on $W_U = p_*W$. This extends to a functor p_* from $\mathfrak{Rep}_B^\infty$ to $\mathfrak{Rep}_T^\infty$.*

In this situation, we have the following adjointness theorem.

Theorem 3.28 (p. 125 of Cartier) [20] *Let (η, Y) be a smooth representation of T, and (σ, W) be a smooth representation of B. Then there is a natural isomorphism:*

$$Hom_B(p_*W, Y) \cong Hom_T(W, p^*Y).$$

This makes p_* a left adjoint to p^*.

[20] P. Cartier. Representations of p-adic groups: a survey. In *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 1*, Proc. Sympos. Pure Math., XXXIII, pages 111–155. Amer. Math. Soc., Providence, R.I., 1979.

Indeed, the coinvariants W_U can be naturally identified with a module obtained by extension of scalars:

$$W_U \cong W \otimes_{\mathbb{C}[B]} \mathbb{C}[T],$$

where $\mathbb{C}[T]$ is viewed as a $\mathbb{C}[B]$-module via the trivial action of U. The result follows from adjointness of ring-extension and pull-back, suitably interpreted.

4 *Representations of GL_2, external theory*

Hereafter, we let $G = \mathbf{GL}_2(k)$, where k is a nonarchimedean local field; very little will be lost by taking $k = \mathbb{Q}_p$. As usual, we study the representations of a complicated group G, by understanding the representations of "easier" subgroups, and the functors of restriction and induction.

In addition, we drop the adjective "smooth" hereafter; all groups will be ℓ-groups, and all representations will be smooth. By "irrep", we mean an irreducible smooth representation.

By the *external theory*, we focus our attention on subgroups H of G which arise as $H = \mathbf{H}(k)$ for *algebraic* subgroups $\mathbf{H} \subset \mathbf{G}$. The primary subgroups of interest are:

$$B = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} : a, d \in k^\times, b \in k \right\},$$

$$T = \left\{ \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} : a, d \in k^\times \right\} \cong k^\times \times k^\times,$$

$$U = \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} : b \in k \right\} \cong k.$$

$$Z = \left\{ \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} : a \in k^\times \right\} = Z(G) \cong k^\times.$$

These subgroups arise as the k-points of algebraic subgroups $\mathbf{B} = \mathbf{TU} \subset \mathbf{G}$. At the level of k-points, one has a semidirect product decomposition $B = T \ltimes U$. We write $p : B \to T$ for the canonical projection.

Representation theory of T

Perhaps this treatment of the torus T is excessive in notation, for such a simple case. The advantage is that everything here generalizes easily to split tori of any rank.

Corresponding to the obvious isomorphism $\mathbf{T} \cong \mathbf{G}_m \times \mathbf{G}_m$, there is an isomorphism of ℓ-groups: $T \cong k^\times \times k^\times$. The algebraic characters and cocharacters of $\mathbf{T}$ are:

$$X^\bullet(\mathbf{T}) = Hom(\mathbf{T}, \mathbf{G}_m) \cong \mathbb{Z}^2,$$

$$X_\bullet = X_\bullet(\mathbf{T}) = Hom(\mathbf{G}_m, \mathbf{T}) \cong \mathbb{Z}^2.$$

There is a canonical perfect pairing:

$$X_\bullet \times X^\bullet \to \mathbb{Z},$$

given by the identification $Hom(\mathbf{G}_m, \mathbf{G}_m) \cong \mathbb{Z}$.

We write $T_\circ$ for the maximal compact subgroup of T; there is a unique maximal compact subgroup, and $T_\circ$ is isomorphic to $\mathcal{O}^\times \times \mathcal{O}^\times$. While this isomorphism is non-canonical, there is a canonical isomorphism:

$$X_\bullet \cong T/T_\circ,$$

given by sending $\alpha \in X_\bullet$ to $\alpha(\varpi) \in T/T_\circ$; the choice of uniformizing element ϖ does not affect the $T_\circ$-coset of $\alpha(\varpi)$. The complex dual torus of T is defined by:

$$\hat{T} = Hom(X_\bullet, \mathbb{C}^\times) = X^\bullet \otimes_{\mathbb{Z}} \mathbb{C}^\times \cong \mathbb{C}^\times \times \mathbb{C}^\times.$$

Since T is abelian, the irreps of $T \cong k^\times \times k^\times$ are one-dimensional – they are given by a pair $\chi = (\chi_1, \chi_2)$ of (smooth) characters

$$\chi_1, \chi_2 : k^\times \to \mathbb{C}^\times.$$

We will pay particular attention to the *unramified* characters of T – these are given by pairs (χ_1, χ_2) of characters, which are both trivial on $\mathcal{O}^\times$. Writing $T_\circ = \mathcal{O}^\times \times \mathcal{O}^\times$, the unramified characters are just $Hom(T/T_\circ, \mathbb{C}^\times)$. Thus the unramified characters of T are described easily by the dual torus:

$$Hom_{unr}(T, \mathbb{C}^\times) = Hom(T/T_\circ, \mathbb{C}^\times) \cong \hat{T} = Hom(\mathbb{Z}, \hat{T}).$$

Much more generally, local class field theory implies that

$$Hom_{cont}(T, \mathbb{C}^\times) \cong Hom_{cont}(W_k, \hat{T}).$$

The unramified characters correspond to those continuous homomorphisms from W_k to $\hat{T}$ that factor through the quotient $W_k^{unr} \cong \mathbb{Z}$. We follow the convention that the unramified character of T corresponding to $t \in \hat{T}$ should correspond to the unramified character of W_k which sends a *geometric* Frobenius element to t.

This is known as the local Langlands corresponence for T, and was generalized by Langlands to arbitrary tori in an article that took thirty years to publish (finally in Pac. J. of Math.[21]).

[21] .

Jacquet functor, supercuspidals

For the classification of irreps of $G = \mathbf{GL}_2(k)$, and more generally in the classification of irreps of reductive p-adic groups, the most important method is parabolic induction and Harish-Chandra's theory of cuspidal representations.[22]

[22] This is the local analogue of the dichotomy between Eisenstein series and cuspforms.

Definition 4.1 *Let (π, V) be a representation of G. The* Jacquet functor *is*

$$J_B^G = p_* \circ Res_B^G : \mathfrak{Rep}_G^\infty \to \mathfrak{Rep}_T^\infty.$$

In particular, $J_B^G V = V_U$ is the space of U-coinvariants of V, viewed as a smooth representation of T.

Definition 4.2 *Let (η, Y) be a representation of T. The* functor of parabolic induction *is*

$$I_B^G = Ind_B^G \circ p^* : \mathfrak{Rep}_T^\infty \to \mathfrak{Rep}_G^\infty.$$

In particular, $I_B^G Y$ consists of uniformly locally constant functions $f : G \to Y$ which satisfy

$$f(tux) = \eta(t)(f(x) \text{ for all } t \in T, u \in U, x \in G,$$

and G acts on this space of functions by right translation.

Theorem 4.3 *The functor J_B^G is left adjoint to I_B^G; for a representation (π, V) of G and a representation (η, Y) of T, there is a natural isomorphism:*

$$Hom_G(V, I_B^G Y) \cong Hom_T(J_B^G V, Y).$$

Proof: Adjointness of Res_B^G and Ind_B^G implies

$$Hom_G(V, I_B^G Y) = Hom_G(V, Ind_B^G p^* Y) \cong Hom_B(Res_B^G V, p^* Y).$$

Adjointness of p^* and p_* implies

$$Hom_B(Res_B^G V, p^* Y) \cong Hom_T(p_* Res_B^G V, Y) = Hom_T(J_B^G V, Y).$$

The naturality of these isomorphisms, i.e., the adjointness of functors, implies the adjointness of I_B^G and J_B^G as required.

Q.E.D

In what follows, it will be more convenient to use the *normalized* parabolic induction and Jacquet functor. Let $\delta : T \to \mathbb{R}_{>0}^\times$ be the character[23] given by:

$$\delta \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} = |a/d|.$$

[23] This is usually called the *modular character*. It describes the effect of T-conjugation on a Haar measure on U. Something like it should be used whenever carrying out induction and restriction involving non-unimodular groups (like B).

Viewing characters of T as pairs of characters of $k^\times$, we find that

$$\delta = (|\cdot|, |\cdot|^{-1}).$$

We write $I_B^G \delta^{1/2}$ for the functor which on objects sends a representation η of T to $I_B^G(\eta \otimes \delta^{1/2})$. Similarly, we write $\delta^{-1/2} J_B^G$ for the functor which sends a representation π of G to $\delta^{-1/2} \otimes J_B^G \pi$.

One advantage of this normalization is that *unitarizability* is preserved; if χ is a unitary character of T (it has values in the unit circle in the complex plane), then there is a natural Hermitian inner product on $I_B^G\delta^{1/2}\chi$; this implies that subrepresentations of $I_B^G\delta^{1/2}\chi$ have complements – it eventually yields complete reducibility of $I_B^G\delta^{1/2}\chi$.

The adjointness of J_B^G and I_B^G implies adjointness of the normalized functors; in particular,

$$Hom_G(V, I_B^G\delta^{1/2}Y) \cong Hom_T(\delta^{-1/2}J_B^G V, Y).$$

The following result makes the representation theory of p-adic groups much easier, in some ways, than the representation theory of real Lie groups:

Proposition 4.4 *The functors I_B^G and J_B^G are exact. Same for the functors $I_B^G\delta^{1/2}$ and $\delta^{-1/2}J_B^G$.*

Proof: (Sketch) Exactness of the functor I_B^G is easy, as is left-exactness of J_B^G. To demonstrate the right-exactness of J_B^G, it suffices to demonstrate the right-exactness of the "U-coinvariant functor" p_*. This follows from the fact that U is the union of compact subgroups – the functor of coinvariants for a compact group is exact (a basic result in group homology with coefficients in a vector space over a characteristic zero field) – and the exactness of direct limits.

For the normalized functors, the result follows by exactness of twisting, which is trivial to check.

Q.E.D

A useful basic result is that I_B^G and J_B^G are compatible with twisting and central characters, in a simple way.

Proposition 4.5 *Let $\chi = (\chi_1, \chi_2)$ be a character of T. Then $I_B^G\chi$ has central character $\chi_1\chi_2$. Furthermore, let χ_0 be a character of $k^\times$ and write $\chi_0\chi$ for the character $(\chi_0\chi_1, \chi_0\chi_2)$ of T; then there is a natural isomorphism of representations of G:*

$$I_B^G(\chi_0\chi) \cong (\chi_0 \circ det) \otimes I_B^G\chi.$$

Proof: The proof is straightforward and left to the reader.

Q.E.D

The Jacquet functor gives an initial classification of irreps of $G = GL_2(k)$:

Definition 4.6 *A representation (π, V) of G is called* supercuspidal *if $J_B^G V = 0$.*

This would not be such an interesting definition if it were not for the following nontrivial theorem, due in various parts, and somewhat independently, to various authors (Bernstein [24], Casselman[25], Adler and Roche[26], among possible others):

Theorem 4.7 *The following conditions are equivalent, for an irrep (π, V) of G, whose central character is $\chi : Z \to \mathbb{C}^\times$:*

1. *(π, V) is supercuspidal – $J_B^G V = 0$.*
2. *For all $v \in V$, and $\lambda \in \tilde{V} = (V')^\infty$, the matrix coefficient $m_{v,\lambda}$ is compactly supported, modulo Z; here $m_{v,\lambda} \in C^\infty(G)$ is defined by*
$$m_{v,\lambda}(g) = \lambda(\pi(g)v).$$
3. *There exists $v \in V$ and $\lambda \in V^\infty$, such that $m_{v,\lambda} \neq 0$ and $m_{v,\lambda}$ is compactly supported, modulo Z.*
4. *(π, V) is injective in the category $\mathfrak{Rep}^\infty_{G,\chi}$.*
5. *(π, V) is projective in the category $\mathfrak{Rep}^\infty_{G,\chi}$.*

In particular, if (π, V) is a smooth representation of G which possesses a central character, there are subrepresentations V^{sc}, V^{ind} such that V^{sc} is supercuspidal, and V^{ind} has no supercuspidal subrepresentation (nor quotient), and $V = V^{sc} \oplus V^{ind}$.

The description of supercuspidal representations of G is beyond the scope of these notes; let us just say that all such representations arise via compact induction, from irreducible representations of compact-modulo-Z subgroups of G, e.g., $Z \cdot GL_2(\mathcal{O})$. We refer to the excellent recent book of Bushnell-Henniart [27] for more.

[24] J. Bernstein. *Represenations of p-adic groups*. Harvard University, 1992. Lectures by Joseph Bernstein. Written by Karl E. Rumelhart.

[25] W. Casselman. *Introduction to the theory of admissible representations of p-adic reductive groups*. 1974. Unpublished manuscript, available online.

[26] Jeffrey D. Adler and Alan Roche. Injectivity, projectivity and supercuspidal representations. *J. London Math. Soc. (2)*, 70(2):356–368, 2004.

[27] Colin J. Bushnell and Guy Henniart. *The local Langlands conjecture for* GL(2), volume 335 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006.

Geometric Decomposition

Consider now an irrep (π, V) of G which is not supercuspidal; that is, $J_B^G V \neq 0$. A priori, $J_B^G V$ is just a smooth represenation of T.

Lemma 4.8 *The representation $J_B^G V$ is finitely-generated as a T-module.*

Proof: Let v be a nonzero vector in V, and let H be an open subgroup of G fixing V. The compactness of $B\backslash G \cong \mathbf{P}^1(k)$ implies that there are a finite number of double cosets in $B\backslash G/H$. Choosing representatives $g_1, \ldots, g_d$ for these cosets, we find that V is generated – as a B-module – by the finite set $\{\pi(g_i)v\}_{1 \leq i \leq d}$. Thus, since U acts trivially on V_U, we find that $V_U = J_B^G V$ is generated – as a T-module – by the projections of the vectors $\pi(g_i)v$ for $1 \leq i \leq d$.

Q.E.D

It follows that

Lemma 4.9 *The representation $J_B^G V$ has an irreducible quotient.*

Proof: The proof is somewhat difficult, and we just sketch the idea. Note that the choice of uniformizing element ϖ yields a decomposition $T \cong X_\bullet \times T_\circ$ of ℓ-groups, where $X_\bullet$ is (non-canonically) isomorphic to $\mathbb{Z}^2$. One may first decompose $J_B^G V$ as a representation of the compact ℓ-group $T_\circ$:

$$J_B^G V = \bigoplus_{\phi \in \hat{T}_\circ} (J_B^G V)_\phi.$$

Each $T_\circ$-isotypic component is then a representation of $X_\bullet \cong \mathbb{Z}^2$. In other words, each $T_\circ$-isotypic component is a $\mathbb{C}[X_\bullet]$-module. Since $J_B^G V$ is nonzero, there exists a $\phi \in \hat{T}_\circ$ such that $(J_B^G)_\phi \neq 0$.

Thus to check that $J_B^G V$ has an irreducible quotient, it suffices to check that $(J_B^G V)_\phi$ has an irreducible quotient as a $\mathbb{C}[X_\bullet]$-module. From our previous study of the representations of $\mathbb{Z}$, it suffices (by Nakayama's lemma) to check that $(J_B^G V)_\phi$ is finitely-generated as a $\mathbb{C}[X_\bullet]$-module. But this follows from the fact that $J_B^G V$ is finitely-generated as a T-module, and $T_\circ$ acts via a character on $(J_B^G V)_\phi$.[28]

[28] In fact, just knowing that $J_B^G V$ is finitely generated as a T-module is enough to show that it has an irreducible quotient, using a Zorn's lemma argument. It is not really necessary to use the $T_\circ$-isotypic components.

Q.E.D

When (π, V) is an irrep of G, we find that $\delta^{-1/2} J_B^G V$ has an irreducible quotient – a character χ of T:

$$Hom_T(\delta^{-1/2} J_B^G V, \chi) \neq 0.$$

We choose to use the normalized functors, for reasons that will become clear. It follows that

$$Hom_G(V, I_B^G \delta^{1/2} \chi) \neq 0,$$

and so V is a subrepresentation of $I_B^G \delta^{1/2} \chi$. Thus the non-supercuspidal representations arise as subrepresentations of *principal series* – representations parabolically induced from characters of tori.

For this reason (and since we are not prepared to discuss supercuspidal representations here), we study the representations $I_B^G \delta^{1/2} \chi$ – the principal series representations of G. The key to studying these representations is the Bruhat decomposition:

$$G = B \sqcup BwB, \quad w = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Here BwB is an open subset of G, and B is its closed complement. The short exact sequence of $\mathbb{C}$-modules:

$$0 \to C_c^\infty(BwB) \to C_c^\infty(G) \to C_c^\infty(B) \to 0$$

is in fact a short exact sequence of smooth representations of $B \times B$ by left and right translation. From this (and a little work to check right-exactness) we obtain a short exact sequence of B-representations (by right-translation):

$$0 \to I^w(\delta^{1/2}\chi) \to I_B^G \delta^{1/2}\chi \to I^1(\delta^{1/2}\chi) \to 0,$$

where:

$$I^w(\chi) = \{f \in C_c^\infty(BwB), \text{ such that } f(tux) = \chi(t)\delta^{1/2}(x)f(x)\},$$

$$I^1(\chi) = \{f : B \to \mathbb{C}, \text{ such that } f(tux) = \chi(t)\delta^{1/2}f(x)\} \cong \mathbb{C}.$$

An explicit and nontrivial[29] computation demonstrates that:

$$\delta^{-1/2} J_B^G(I^w(\delta^{1/2}\chi)) \cong \chi^w, \quad \delta^{-1/2} J_B^G(I^1(\delta^{1/2}\chi)) \cong \chi.$$

To summarize, there is a short exact sequence of T-representations

$$0 \to \chi^w \to \delta^{-1/2} J_B^G I_B^G \delta^{1/2}\chi \to \chi \to 0. \tag{1}$$

Here,

$$\chi = (\chi_1, \chi_2), \quad \chi^w = (\chi_2, \chi_1).$$

If χ and χ^w are distinct characters of T, then the short exact sequence splits and:

$$\delta^{-1/2} J_B^G I_B^G \delta^{1/2}\chi \cong \chi \oplus \chi^w.$$

Lemma 4.10 *If W is any subquotient of $I_B^G \delta^{1/2}\chi$, then $J_B^G W \neq 0$.*

Proof: If $J_B^G W = 0$, then W is supercuspidal. It follows, from injectivity and projectivity of supercuspidals[30], and the fact that $I_B^G \delta^{1/2}\chi$ has a central character $\delta^{1/2}\chi_1\chi_2$, that the subquotient W of $I_B^G \delta^{1/2}\chi$ also arises as a submodule. Hence

$$Hom(W, I_B^G \delta^{1/2}\chi) \neq 0.$$

By adjointness,

$$Hom(J_B^G W, \delta^{1/2}\chi) \neq 0.$$

This contradicts the fact that W is supercuspidal.

Q.E.D

Corollary 4.11 *The representation $I_B^G \delta^{1/2}\chi$ has length at most two.*

Proof: The exactness of the functor J_B^G, the previous lemma, and the fact that $\delta^{-1/2} J_B^G I_B^G \delta^{1/2}\chi$ is two-dimensional implies this corollary.

[29] A geometric argument – that $\mathbf{B}w\mathbf{B}$ is isomorphic to $\mathbf{B} \times \mathbf{U}$ as a k-variety – implies that $I^w(\delta^{1/2}\chi)$ is one-dimensional. Seeing that $I^1(\delta^{1/2}\chi)$ is one-dimensional is easier. One identifies the projection of $I^w(\delta^{1/2}\chi)$ onto its U-coinvariants with an integral over U – tracking through the T-action proves the result.

[30] Really, it is deceptive to utilize injectivity and projectivity of supercuspidals for this sort of result. The proof of injectivity and projectivity of supercuspidals relies on results like this lemma, to my recollection. It is much better to prove this lemma using Jacquet's lemma, and compact subgroups with Iwahori decomposition.

Q.E.D

The precise conditions for reducibility of $I_B^G \delta^{1/2} \chi$ are given by the following

Theorem 4.12 *The representation $I_B^G \delta^{1/2} \chi$ is reducible if and only if*

$$\chi_1 = |\cdot| \chi_2, \text{ or } \chi_1 = |\cdot|^{-1} \chi_2.$$

Equivalently, $I_B^G \delta^{1/2} \chi$ is reducible if and only if

$$\chi = \delta^{\pm 1} \chi^w.$$

This theorem requires a lot of work – we refer to the exposition of Tadic for a nice treatment. Partial results follow from Frobenius reciprocity and the short exact sequence (**??**): we find that

$$End_G(I_B^G \delta^{1/2} \chi) \cong Hom_T(\delta^{-1/2} J_B^G I_B^G \delta^{1/2} \chi, \chi).$$

We find two cases:

1. The space $End_G(I_B^G \delta^{1/2} \chi)$ is one-dimensional, if $\chi \neq \chi^w$, or if $\chi = \chi^w$ and the extension $\delta^{-1/2} J_B^G I_B^G \delta^{1/2} \chi$ of χ by itself is nontrivial.
2. The space $End_G(I_B^G \delta^{1/2} \chi)$ is two-dimensional if $\chi = \chi^w$ and the extension $\delta^{-1/2} J_B^G I_B^G \delta^{1/2} \chi$ of χ by itself splits.

By Schur's lemma, if $End_G(I_B^G \delta^{1/2} \chi)$ is two-dimensional, then $I_B^G \delta^{1/2} \chi$ is reducible; but the above observation implies that $\chi = \chi^w$, and Theorem **??** implies that there is no reducibility when $\chi = \chi^w$ (only when $\chi = \delta^{\pm 1} \chi^w$). Hence we find that

Corollary 4.13 *The representation $I_B^G \delta^{1/2} \chi$ is either irreducible, or else is a nonsplit extension of one irreducible representation of G by another irreducible representation of G.*

Proof: If $I_B^G \delta^{1/2} \chi$ is reducible, we find that its G-endomorphisms form a one-dimensional space. Hence it cannot be decomposed into the direct sum of irreducible representations. Since it has length at most two, the result follows immediately.

Q.E.D

One example is particularly easy to see, and important for applications:

Example 4.14 *Considering $\chi = \delta^{-1/2}$, we find that $I_B^G \delta^{1/2} \chi = I_B^G \mathbb{C}$ is a reducible representation of G, of length two. There is a short exact sequence of smooth representations of G:*

$$0 \to \mathbb{C} \to I_B^G \mathbb{C} \to St \to 0.$$

The embedding of $\mathbb{C}$ into $I_B^G\mathbb{C}$ takes a complex number to the corresponding constant function on G. Since its image is clearly one-dimensional, and $I_B^G\mathbb{C}$ is infinite-dimensional, there must be a nontrivial quotient. This quotient is called the Steinberg *representation.*

The symmetry between χ_1 and χ_2 manifests in a rational family (rational, in the parameter $\chi \in Hom_{cont}(T, \mathbb{C}^\times)$) of intertwining operators, from $I_B^G\delta^{1/2}\chi$ to $I_B^G\delta^{1/2}\chi^w$.

Proposition 4.15 *Suppose that $\chi \neq \chi^w$. Then $I_B^G\delta^{1/2}\chi$ is isomorphic to $I_B^G\delta^{1/2}\chi^w$.*

Proof: By Frobenius reciprocity, there is a natural $\mathbb{C}$-linear isomorphism

$$Hom_G(I_B^G\delta^{1/2}\chi, I_B^G\delta^{1/2}\chi^w) \cong Hom_T(\delta^{-1/2}J_B^G I_B^G\delta^{1/2}\chi, \chi^w).$$

Recall the short exact sequence of representations of T **??**:

$$0 \to \chi^w \to \delta^{-1/2}J_B^G I_B^G\delta^{1/2}\chi \to \chi \to 0.$$

It follows that if $\chi \neq \chi^w$, then the above sequence splits, $I_B^G\chi$ and $I_B^G\chi^w$ are irreducible, and hence are isomorphic to each other.

Q.E.D

In fact, the intertwining operators, which exist by Frobenius reciprocity, form a complex algebraic family over (a Zariski-dense subset of) the variety $Hom_{cont}(T, \mathbb{C}^\times)$. However, these operators have zeros and poles, which correspond to the reducibility points of the principal series representations.

Unramified principal series

Especially important for global applications are the unramified principal series; these are the representations $I_B^G\delta^{1/2}\chi$, when $\chi : T/T_\circ \to \mathbb{C}^\times$ is an unramified character of T. In particular,

$$\chi = (\chi_1, \chi_2), \quad \chi_i(x) = s_i^{val(x)},$$

for some nonzero complex numbers s_1, s_2. The pair (s_1, s_2) can be thought of as an element of $\hat{T}$, if one wishes to be canonical. For simplicity, we define

$$I(s_1, s_2) = I_B^G\delta^{1/2}\chi, \text{ when } \chi\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} = (s_1^{val(a)}, s_2^{val(d}).$$

From Proposition **??**, when $s_1 \neq s_2$, there is an isomorphism:

$$I(s_1, s_2) \cong I(s_2, s_1).$$

We find a reducibility point when $\chi_1 = |\cdot|^{\pm 1}\chi_2$. In other words,

Proposition 4.16 *The unramified principal series $I(s_1, s_2)$ is reducible if and only if $s_1 = q^{\pm 1} s_2$. Here, we recall that $q = \#(\mathcal{O}/\varpi)$ is the order of the residue field of k.*

PROOF: This follows from the previous result on reducibility of principal series representations, and changing notation.

Q.E.D

If $s_1 = q^{-1} s_2$, then we find that

$$\delta^{1/2} \cdot (\chi_1, \chi_2) = (|\cdot|^{1/2}\chi_1, |\cdot|^{-1/2}\chi_2) = (|\cdot|^{-1/2}\chi_2, |\cdot|^{-1/2}\chi_2).$$

It follows that

$$I_B^G \delta^{1/2}\chi \cong |det|^{-1/2} s_2^{val(det)} \otimes I_B^G \mathbb{C}.$$

In this case, $I_B^G \delta^{1/2}\chi$ has an irreducible subrepresentation and irreducible quotient:

$$0 \to |\cdot|^{-1/2} s^{val(det)} \to I(s_1, s_2) \to |\cdot|^{-1/2} s^{val(det)} \otimes St \to 0.$$

If $s_1 = q s_2$, then one finds a similar short exact sequence, with a twisted trivial representation as a quotient, and twisted Steinberg representation as a subrepresentation.

To summarize, we have a two-dimensional complex algebraic variety[31] $\hat{T} = MSpec(\mathbb{C}[s_1^{\pm 1}, s_2^{\pm 1}])$, acted upon by a finite group $W = \{1, w\}$, where w switches s_1 and s_2. There's a W-stable subvariety $\hat{T}_{red}$ cut out by the equations $s_1 = q^{\pm 1} s_2$.

There is a complex algebraic family (see Bernstein[32] for the precise meaning) of representations $I(s_1, s_2)$ of G, parameterized by $(s_1, s_2) \in \hat{T}$, which is generically irreducible, and everywhere satisfies the conclusion of Schur's lemma. The group $W = \{1, w\}$ acts on $\hat{T}$, and on the Zariski-open irreducible locus $\hat{T} - \hat{T}_{red}$. Intertwining operators make this complex algebraic family of representations into a W-equivariant sheaf, when pulled back to $\hat{T} - \hat{T}_{red}$.

In any case, we find that the irreducible constituents of unramified principal series representations are parameterized by the following data:

1. An unordered pair $\{s_1, s_2\}$ of nonzero complex numbers, such that $s_1 \neq q^{\pm 1} s_2$ or...
2. An ordered pair (s_1, s_2) of nonzero complex numbers, such that $s_1 = q^{-1} s_2$ and an additional "bit of information" encoding whether one takes the twisted trivial subrepresentation or twisted Steinberg quotient representation.

To such data, we associate the following Langlands parameters:

[31] We identify complex algebraic varieties with their $\mathbb{C}$-points here.

[32] J. N. Bernstein. Le "centre" de Bernstein. In *Representations of reductive groups over a local field*, Travaux en Cours, pages 1–32. Hermann, Paris, 1984. Edited by P. Deligne.

1. The $GL_2(\mathbb{C})$-conjugacy class containing the semisimple element $\begin{pmatrix} s_1 & 0 \\ 0 & s_2 \end{pmatrix}$. (Note that this only depends on, and uniquely determines, the unordered pair $\{s_1, s_2\}$ of nonzero complex numbers.
2. The $GL_2(\mathbb{C})$-conjugacy class of the pair (t, N), where t is the semisimple element $\begin{pmatrix} s_1 & 0 \\ 0 & s_2 \end{pmatrix}$ (with $s_1 = q^{-1}s_2$), and N is a nilpotent element of $M_2(\mathbb{C})$ satisfying $tNt^{-1} = qN$; for any such (s_1, s_2), there are two such conjugacy classes of pairs: one contains $(t, 0)$ and the other contains (t, N) with $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

The first case can also be thought of as a conjugacy class of pairs (t, N) with $tNt^{-1} = qN$; but when $s_1 \neq q^{\pm 1}s_2$, the only nilpotent N satisfying that identity is zero. In the second case, the extra "bit of information" given by whether $N = 0$ or $N \neq 0$ corresponds to the extra "bit of information" given by whether one chooses the twisted trivial representation of the twisted Steinberg representation, respectively.

If an irreducible constituent of an unramified principal series representation (π, V) corresponds to a parameter (t, N) as above (t semisimple in $GL_2(\mathbb{C})$ and N nilpotent in $M_2(\mathbb{C})$), then the standard (degree 2) L-function of (π, V) is:

$$L(\pi, Stand) = det(1 - tX|Ker(N)).$$

5 *Representations of GL_2, internal theory*

Let K be an open compact subgroup of $G = \mathbf{GL}_2(k)$. It is important to study representations with K-fixed vectors; in order to have a good *category* of representations, we define $\mathfrak{Rep}_G^K$ to be the category of smooth representations of G which are generated (as G-represenations) by their K-fixed vectors. These are called K-spherical representations. For general K, this category is not stable under subquotients!

Let $H(G, K)$ be the Hecke algebra of compactly supported, K-bi-invariant functions on G:

$$H(G, K) = C_c^\infty(K\backslash G/K).$$

If (π, V) is a K-spherical representation, then V^K is naturally an $H(G, K)$-module, via

$$\pi(f)v = \int_G f(g)\pi(g)v dg.$$

If $f_1, f_2 \in H(G,K)$, then

$$\pi(f_1)\pi(f_2)v = \pi(f_1 * f_2)v,$$

where the convolution is defined by

$$[f_1 * f_2](g) = \int_G f_1(h) f_2(h^{-1}g) dh.$$

In fact, this gives an equivalence of categories, from the category of modules over the *convolution* algebra $H(G,K)$ and the category of K-spherical representations.

These categories are somewhat mysterious in general, but when $K = \mathbf{GL}_2(\mathcal{O})$, we have the category of *unramified representations*. These are well-understood; moreover in the factorization of automorphic representations, irreducible unramified representations occur for almost all primes.

Unramified representations

Hereafter, let $K = \mathbf{GL}_2(\mathcal{O})$. The remarkable theorem about the spherical Hecke algebra is the following:

Theorem 5.1 *Define, for $f \in H(G,K)$, the* Satake transform $Sf \in C_c^\infty(T)$

$$[Sf](t) = \delta(t)^{-1/2} \int_U f(ut) du = \delta(t)^{1/2} \int_U f(tu) du.$$

Then

$$Sf \in H(T, T_\circ) = C_c^\infty(T/T_\circ)^W = \mathbb{C}[X_\bullet]^W,$$

where $W = \{1, w\}$. Moreover, S determines an isomorphism of algebras:

$$H(G,K) \cong \mathbb{C}[X_\bullet]^W.$$

In particular, this theorem implies that $H(G,K)$ is a *commutative* $\mathbb{C}$-algebra! Highest weight theory, for the algebraic representations of $GL_2(\mathbb{C})$, implies that

$$\mathbb{C}[X_\bullet(\mathbf{T})]^W = \mathbb{C}[X^\bullet(\hat{\mathbf{T}})]^W \cong Rep(GL_2(\mathbb{C})),$$

where $Rep(GL_2(\mathbb{C}))$ is the complexification of K_0 of the category of finite-dimensional algebraic representations of $GL_2(\mathbb{C})$ – i.e., the complexified representation ring of $GL_2(\mathbb{C})$.

As the category of spherical representations of G is equivalent to the category of $H(G,K)$-modules, which is equivalent to the category of $\mathbb{C}[X_1^{\pm 1}, X_2^{\pm 1}]^W$-modules. It follows that an *irreducible* unramified representation of G is one-dimensional – determined by two nonzero complex numbers (s_1, s_2), modulo switching;

there is a natural bijection between the isomorphism classes of irreducible unramified representations of G and *unordered pairs* $\{s_1, s_2\}$ of nonzero complex numbers.

More canonically, there is a natural bijection between the set of isomorphism classes of irreducible unramified representations of G and the W-orbits on $\hat{T}$.

Connection to unramified principal series

Let $I \subset K$ be the *Iwahori subgroup*, consisting of matrices in $\mathbf{GL}_2(\mathcal{O})$ whose lower-left entry is in $\varpi\mathcal{O}$. Recall that $T_\circ = \mathbf{T}(\mathcal{O}) \cong \mathcal{O}^\times \times \mathcal{O}^\times$. The following is a fundamental theorem of Borel and Matsumoto:

Theorem 5.2 *Let (π, V) be an admissible (smooth and finite-length certainly suffices) representation of G. Consider the natural projection map $V \to V_U$ from V onto the space of $J_B^G V$. This projection map induces an isomorphism of complex vector spaces:*

$$V^I \to (V_U)^{T_\circ}.$$

A corollary of this result is the following:

Corollary 5.3 *If (π, V) is a K-spherical admissible representation of G, then $J_B^G V \neq 0$. If moreover, (π, V) is an irreducible unramified representation of G, then $J_B^G V$ has an unramified character of T as a subquotient.*

By adjointness, and what we know about unramified principal series, we find that

Corollary 5.4 *If (π, V) is an irreducible unramified representation of G, then (π, V) occurs as a subquotient in an unramified principal series representation $I_B^G \delta^{1/2} \chi$, where $\chi : T/T_\circ \to \mathbf{C}^\times$ is uniquely determined by V up to the action of W.*

From this result, we find that an irreducible unramified representation (π, V) of G yields two pairs of complex numbers:

1. Since (π, V) is associated to an irreducible $H(G, K)$-module, we obtain two "Hecke eigenvalues" s_1, s_2 (up to switching). These are called the Satake parameters of (π, V), since they arise from the Satake isomorphism from $H(G, K)$ to $H(T, T_\circ)$.

2. Since (π, V) occurs in an unramified principal series representation, we find that (π, V) is a subquotient of $I(t_1, t_2)$, for nonzero complex numbers t_1, t_2, uniquely determined, up to switching.

Furthermore, although the unramified principal series representation $I_B^G \delta^{1/2} \chi$ may be reducible, it has a unique unramified subrepresentation – the twisted trivial representation (with unramified twist, of course) is always unramified, and the twisted Steinberg representation is never unramified (has no K-fixed vectors).

The connection between these is the following significant theorem:

Theorem 5.5 *The unordered pair $\{s_1, s_2\}$ equals the unordered pair $\{t_1, t_2\}$.*

Let π_{s_1,s_2} denote the irreducible spherical representation of $GL_2(k)$ with parameters $s_1, s_2 \in \mathbb{C}^\times$.

The impact of this theorem, for the theory of modular forms, is the following: Let f be a classical modular form for a congruence subgroup $\Gamma_0(N)$; suppose that f is a cuspidal newform, of some Nebentypus, for good measure. Then one associates to f an automorphic representation $\Pi = \bigotimes' \pi_v$, where the (restricted) tensor product is over all places v of $\mathbb{Q}$. At all primes p not dividing N, the representation π_p is irreducible and unramified.

The previous theorem tells us that the eigenvalue of the T_p operator (and the Nebentypus character), which determines the Hecke eigenvalue and hence the Satake parameter for the representation π_p, also determines the isomorphism class of the representation π_p. The representation π_p is precisely the irreducible unramified constituent of the unramified principal series $I(s_1, s_2)$, where (s_1, s_2) is the Satake parameter deduced from the Hecke eigenvalue of T_p.

Slightly more generally, if p divides N, but p^2 does not divide N, the representation π_p ends up being isomorphic to a twist of the Steinberg representation; proving this requires some analysis of the Iwahori Hecke algebra $H(G, I)$ instead of $H(G, K)$.

References

Jeffrey D. Adler and Alan Roche. Injectivity, projectivity and supercuspidal representations. *J. London Math. Soc. (2)*, 70(2):356–368, 2004.

J. Bernstein. *Represenations of p-adic groups*. Harvard University, 1992. Lectures by Joseph Bernstein. Written by Karl E. Rumelhart.

J. N. Bernstein. Le "centre" de Bernstein. In *Representations of reductive groups over a local field*, Travaux en Cours, pages 1–32. Hermann, Paris, 1984. Edited by P. Deligne.

Émile Borel. *Sur quelques points de la théorie des fonctions*. Paris., 1894. Original from Columbia University.

Colin J. Bushnell and Guy Henniart. *The local Langlands conjecture for* GL(2), volume 335 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006.

P. Cartier. Representations of p-adic groups: a survey. In *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 1*, Proc. Sympos. Pure Math., XXXIII, pages 111–155. Amer. Math. Soc., Providence, R.I., 1979.

W. Casselman. *Introduction to the theory of admissible representations of p-adic reductive groups*. 1974. Unpublished manuscript, available online.

Brian Conrad. Weil and Grothendieck approaches to adelic points. Unpublished notes, available online.

S. DeBacker. *Some notes on the representation theory of reductive p-adic groups*. Available online.

P. Garrett. Smooth representations of totally disconnected groups. Introductory notes, available online. Updated July 8, 2005.

Hervé Jacquet. Sur les représentations des groupes réductifs p-adiques. *C. R. Acad. Sci. Paris Sér. A-B*, 280:Aii, A1271–A1272, 1975.

Vladimir Platonov and Andrei Rapinchuk. *Algebraic groups and number theory*, volume 139 of *Pure and Applied Mathematics*. Academic Press Inc., Boston, MA, 1994. Translated from the 1991 Russian original by Rachel Rowen.

MODULAR FORMS AND AUTOMORPHIC REPRESENTATIONS

DENIS TROTABAS

Contents

1. Some notations.

Let $\mathfrak{H} = \{z \in \mathbf{C}; \Im(z) > 0\}$ be the Poincaré upper-half plane.

Let k and N be two integers, and, as usual, $\Gamma_0(N)$ be the subgroup of $\mathrm{SL}(2, \mathbf{Z})$ of matrices whose lower left entries are divisible by N. It acts on $\mathfrak{H}$ by fractional linear transformations: $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \cdot z = \frac{az+b}{cz+d}$.

Let χ be a Dirichlet character modulo q: it defines a character on $\Gamma_0(N)$, by evaluating χ at the upper left entry. It will be convenient to define $\chi(n) = 0$ if the integer n is not coprime with N.

If X is a finite set, $|X|$ denotes its cardinality; we reserve the letters p, ℓ for prime numbers, and n, m for integers.

The letters K, E, k (resp. K_λ) denote fields (resp. the completion of K with respect to the valuation associated to λ), and $\mathcal{O}_K, \mathcal{O}_\lambda$ stand for the rings of integers of K, K_λ in the relevant situations.

The set of adeles of $\mathbf{Q}$ is denoted $\mathbb{A}_{\mathbf{Q}}$, and for a finite set of primes S containing ∞, one denotes $\mathbb{A}_{\mathbf{Q},S} = \prod_{v \in S} \mathbf{Q}_v \times \prod_{v \notin S} \mathbf{Z}_v$. The finite adeles are denoted $\mathbb{A}_f$.

For a complex number z, the notation $\mathrm{e}(z)$ stands for $\exp(2\pi i z)$.

The notation $f(x, A) \ll_A g(x)$ means that for any A, there exists a real number $C(A)$ such that for any x, $|f(x, A)| \leq C(A) \cdot |g(x)|$; if one adds "as

$x \to \infty$", it means that the last inequality holds for $x \geq x(A)$ for some real number $x(A)$. In the same spirit, the notation $f(x) = o_{x\to x_0}(g(x))$ (resp. $f(x) = O_{x\to x_0}(g(x))$) means that the quotient $f(x)/g(x)$ is defined in a (pointed) neighbourhood of x_0, and that $|f(x)/g(x)|$ tends to zero (resp. stays bounded) when x tends to x_0.

Some spaces of funtions: let $\mathfrak{X}$ be a locally compact Hausdorff space.

- $\mathcal{C}_c(\mathfrak{X})$ is the space of continuous compactly supported complex valued funcions.
- $\mathcal{C}_c^\infty(\mathfrak{X})$ denotes the subspace of smooth functions in the latter (when $\mathfrak{X}$ is a manifold, this means "locally constant" if the manifold is totally disconnected).

2. Modular forms

2.1. For any holomorphic function f defined on $\mathfrak{H}$ and $\gamma \in \Gamma_0(N)$, we define:

$$f_{|\gamma}(z) = \chi(\gamma)^{-1}(cz+d)^{-k} f(\gamma \cdot z)$$

Consider the following properties:

(M1): For any $\gamma \in \Gamma_0(N)$, $f_{|\gamma} = f$. This implies, by Fourier analysis, that for any $\sigma \in \mathrm{SL}(2,\mathbf{Z})$, there exists a positive integer $h(\sigma)$ (with $h(\mathrm{I}) = 1$) such that one has an absolutely convergent decomposition:

$$f_{|\gamma}(z) = \sum_{n\in\mathbf{Z}} c_n(f,\sigma)\mathrm{e}(nz/h(\sigma))$$

The holomorphy at $i\infty$ is then expressed by:

(M2): For any $\sigma \in \mathrm{SL}(2,\mathbf{Z})$, $c_n(f,\sigma) = 0$ for all negative n.

"Cuspidality" is:

(M2′): For any $\sigma \in \mathrm{SL}(2,\mathbf{Z})$, $c_n(f,\sigma) = 0$ for all $n \geq 0$.

2.2. The space of modular forms of weight k, level q and nebentypus χ is the set of holomorphic functions satisfying (M1) and (M2) above; the subspace of modular forms satisfying (M2′) as well is called the space of cusp forms, noted $\mathcal{S}_k(N,\chi)$. The latter is finite dimensional (as is the first), and equipped with the Petersson inner product, invariant under the group action (it is a quotient of a Haar measure on $\mathfrak{H} = \mathrm{SO}_2(\mathbf{R})\backslash\mathrm{SL}_2(\mathbf{R})$):

$$\langle f,g\rangle = \int_{\Gamma_0(N)\backslash\mathfrak{H}} f(x+iy)\overline{g(x+iy)}y^k\frac{dxdy}{y^2}$$

Note right now that by taking $\gamma = -\mathrm{I}$, (M1) gives $f(z) = (-1)^k\chi(-1)f(z)$, so if χ and k don't have the same parity, the space of modular forms is $\{0\}$; we shall exclude this case.

2.3. **Hecke operators.** On the space of modular forms of weight k and level q, one has the so-called Hecke operators, defined as follows. Let $n \geq 1$ be an integer, and let $\Delta_0(N) = \{\gamma = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{M}_2(\mathbf{Z}) \ : \ \det(\gamma) > 0, N|c, (a, N) = 1\}$. For $\alpha \in \Delta_0(N)$, one defines first:

$$T_\alpha(f)(z) = \det(\alpha)^{k-1}\chi(\alpha)^{-1}(cz+d)^{-k} f\left(\frac{az+b}{cz+d}\right)$$

To define the level n Hecke operator, one considers the set $\{\alpha \in \Delta_0(N) \ : \ \det(\alpha) = n\}$ on which $\Gamma_0(N)$ acts on the left. One proves that one can write it as a finite disjoint union $\sqcup_j \Gamma_0(N)\alpha_j$, and one defines:

$$T_n(f)(z) = \sum_j (T_{\alpha_j} f)(z)$$

More explicily, one has:

$$\{\alpha \in \Delta_0(N) \ : \ \det(\alpha) = n\} = \bigcup_{\substack{ad=n \\ a>0 \\ (a,q)=1}} \bigcup_{0 \leq b \leq d-1} \Gamma_0(N) \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$$

from which one deduces ($\chi(a) = 0$ if a and N are not coprime):

$$T_n(f)(z) := n^{k-1} \sum_{\substack{ad=n \\ a>0 \\ 0 \leq b \leq d-1}} \chi(a) d^{-k} f\left(\frac{az+b}{d}\right)$$

With this definition, one sees easily that the T_n's preserve the modularity and cupsidality. One can then give the action of the T_p, for p prime, on the Fourier expansion of a modular form (but the modularity is hardly seen from this expression):

- If $(p, N) = 1$, $T_p(f)(z) = \sum_n c_{pn}(f)\mathrm{e}(nz) + \chi(p)p^{k-1}\sum_n c_n(f)\mathrm{e}(pnz)$ – p is called a good prime.
- If $p|N$, $T_p(f)(z) = \sum_n c_{pn}(f)\mathrm{e}(nz)$ – p is a bad prime.

The Hecke operators preserve the space of cusp forms; the Hecke operators at good primes all commute, and are normal with respect to the Petersson inner product. These important facts are explained in Miyake [M], as are the multiplicativity relations. In particular, if f is an eigenfunction for all the Hecke operators at good primes, with eigenvalues $\{a_p(f)\}$, one has $c_p(f) = c_f(1)a_p(f)$ at good p. To diagonalize further the Hecke operators, and get a good definition of L-series, it is necessary to introduce

2.4. **Newforms and oldforms.** Suppose χ defines a Dirichlet character modulo N', for $N'|N$. For any cusp form g in $\mathcal{S}_k(N', \chi)$, one checks easily that $z \mapsto g(dz)$ defines an element of $\mathcal{S}_k(N, \chi)$, for any $d|(N/N')$. Let

$$\mathcal{S}_k^{\mathrm{old}}(N, \chi) = \bigcup_{\substack{\chi \text{ factors through } N'|N \\ d|(N/N')}} \{z \mapsto g(dz) \ : \ g \in \mathcal{S}_k(N', \chi)\}$$

be the space of oldforms, and let

$$\mathcal{S}_k^{\text{new}}(N,\chi) = \mathcal{S}_k^{\text{old}}(N,\chi)^{\perp}$$

be the space of newforms (it may be zero!). Then it can be shown that the whole Hecke algebra (i.e. including bad primes) can be diagonalized on the space of newforms. The primitive Hecke eigenforms (those with $c_1(f) = 1$) have pairwise distinct systems of eigenvalues outside a finite number of primes ("multiplicity one", well explained in the adelic setting by Casselman [C], cf. Gelbart [G] as well). Their L-series have an Euler product, absolutely convergent if $\Re(s) > 1 + k/2$:

$$L(s,f) := \sum_n \frac{a_n(f)}{n^s} = \prod_p L(s,f_p)$$

with

$$\begin{aligned} L(s,f_p) &= \left(1 - a_p(f)p^{-s} + \chi(p)p^{k-1-2s}\right)^{-1} \\ &= \left(1 - \alpha_1(p,f)p^{-s}\right)^{-1}\left(1 - \alpha_2(p,f)p^{-s}\right)^{-1} \end{aligned}$$

at a good prime p, and

$$L(s,f_p) = \left(1 - a_p(f)p^{-s}\right)^{-1}$$

at a bad prime, along with an analytic continuation (easy to see with the Mellin transform), functional equation – cf. Bump [Bu], Miyake [M], Iwaniec [I], etc.

When one proves a theorem, one can often reduce it to the case of newforms, thanks to this decomposition.

2.5. **Ramanujan conjecture.** Let f be a primitive newform. The Ramanujan conjecture is the following inequality:

$$|a_p(f)| \leq 2p^{\frac{k-1}{2}}$$

for good p, which is equivalent to $|\alpha_i(p,f)| = p^{\frac{k-1}{2}}$. It has been a theorem for 35 years now, proven by Deligne for weight greater than two. In the case of bad p one can compute the possibilities for $a_p(f)$ rather explicitly (see [M]).

2.6. **Rationality properties.** Let $f \in \mathcal{S}_k(N,\chi)$ be a eigenform for all the Hecke operators at good primes, with Hecke eigenvalues $\{a_f(p)\}_{p \not| N}$. Then:

$$\mathbf{Q}(f) := \mathbf{Q}(a_f(p), \chi(p) : p \not| N)$$

is a finite extension of $\mathbf{Q}$, and all the Hecke eigenvalues are integers in this extension. If the nebentypus is trivial, then this extension in totally real. Serre explains all of this in terms of arithmetic geometry in his Durham lectures.

2.7. An interesting problem is the evaluation of the dimension of the space of cusp forms, when one or more of the parameters (k, N) vary. For instance, using Eichler-Selberg trace formula one can prove that (see Knightly-Li [KL] theorem 29.5):

$$\dim(\mathcal{S}_k(N,\chi)) = \frac{k-1}{12}\psi(N) + O\Big(N^{1/2}\tau(N)\Big) \tag{1}$$

uniform in $k \geq 2$ and N, where $\psi(N) = q\prod_{p|N}(1+p^{-1})$ and $\tau(N)$ is the number of divisors of N.

Similarly, one can bound the dimension of the space of newforms using the Petersson trace formula (Iwaniec-Luo-Sarnak [ILS]), and one has a uniform estimate for N squarefree, $k \geq 2$:

$$\dim(\mathcal{S}_k^{\text{new}}(N,\chi)) = \frac{k-1}{12}\varphi(N) + O\Big((kN)^{2/3}\Big) \tag{2}$$

with $\varphi(N) = q\prod_{p|N}(1-p^{-1})$ the Euler phi function.

3. Representation theory

If G is a locally compact group, V a complex topological vector space ($\mathbf{C}$ is endowed either with the discrete or the euclidean topology), a *representation* of G in V is a group homomorphism $\rho : G \to \mathrm{Aut}(V)$, such that the mapping $(g, v) \in G \times V \mapsto \rho(g)v \in V$ is continuous. One says: "(ρ, V) is a representation", or simply "let ρ be a representation". But this notion is not sufficient in applications: if G is an algebraic group (over $\mathbf{Q}$ say), then $G(\mathbf{R})$ has a natural structure of a Lie group in which case the notion of $(\mathfrak{g}, K)$-module is important. On $G(\mathbf{Q}_p)$ for $p \geq 2$, one is led to consider also "smooth" representations. On $G(\mathbb{A}_{\mathbf{Q}})$, the notion of "automorphic representation" has at least three interpretations. The point of this section is to provide some background and references on this topic. We chose to minimize the amount of references, but all that follows can be found in any serious book on the subject.

3.1. Let $\mathfrak{H}$ be a Hilbert space. A *unitary* representation is a representation $(\rho, \mathfrak{H})$ such that $\rho(g)$ is unitary for any g in G. Examples:

(1) Given $f \in L^2(G)$ (dg here is a right Haar measure), put

$$R(g)(f)(h) := f(hg)$$

Then $(L^2(G), R)$ is a unitary representation of G, called the right regular representation. Indeed, let $f_1, f_2 \in L^2(G)$, $x, y \in G$. As $\|R(x)f_1 - R(y)f_2\|_2 \leq \|f_1 - f_2\|_2 + \|R(xy^{-1})f_2 - f_2\|_2$, it suffices to prove that for any $f \in L^2(G)$:

$$\lim_{x\to e} \|R(x)f - f\|_2 = 0$$

By using exactly the same argument, and approximating f with compactly supported continuous φ, it suffices to do it for φ instead of f, and this is trivial.

One can extend this idea to the general situation, and prove using the uniform boundedness theorem that if for any v in a fixed dense subset of $\mathfrak{H}$ and $w \in \mathfrak{H}$, the mapping $g \mapsto \langle \rho(g)v, w\rangle$ is continuous then ρ is a representation: see [Wal] lemma 1.1.3, [War] proposition 4.2.2.1, [Ro] chapter 13.

(2) With exactly the same proof, the right action of G on $L^2(H\backslash G)$ (H is a closed subgroup of G, both of which are unimodular say) provides a representation.

(3) If G is compact, and $(\rho, \mathfrak{H})$ a representation, one can show that one can put an inner product on $\mathfrak{H}$, without changing the topology, so that ρ becomes unitary: see lemma 1.4.8 of [Wal] (the idea is to average over G the inner product of course).

One says that $(\rho, \mathfrak{H})$ is *irreducible* if $\mathfrak{H}$ has no closed nontrivial G-invariant proper subspaces. When a representation is not irreducible, it may (or may not) be a Hilbert sum of irreducible subrepresentations. Two unitary representations $(\rho, \mathfrak{H}), (\rho', \mathfrak{H}')$ are *equivalent* if there exist a G-equivariant linear homeomorphism between $\mathfrak{H}$ and $\mathfrak{H}'$: it can be shown that such an isomorphism can be chosen to be an isometry (cf. [Bo2], 5.2). Note that if $(\rho, \mathfrak{H})$ is an irreducible unitary representation of G, then $\mathrm{span}(\rho(g)v \ : \ g \in G)$ is dense in $\mathfrak{H}$. This implies that the Hilbert dimension of $\mathfrak{H}$ is less than $\mathrm{card}(G)$, and therefore the *set* of unitary irreducible representations of G up to equivalence (or isomorphism) is a well defined object: it is denoted $\widehat{G}$.

Theorem 3.1 (Schur's lemma). *If $(\rho, \mathfrak{H})$ is irreducible, then* $\mathrm{Hom}_G(\mathfrak{H}, \mathfrak{H}) = \mathbf{C}\mathrm{Id}_{\mathfrak{H}}$. *Furthermore, if $(\rho', \mathfrak{H}')$ is another (not necessarily irreducible) unitary representation, then any nonzero element of* $\mathrm{Hom}_G(\mathfrak{H}, \mathfrak{H}')$ *is a positive real scalar multiple of an isometry.*

Reference: [Wal] section 1.2, [KL] proposition 10.14.

Application: Let Z denote the center of G, and let $(\pi, \mathfrak{H})$ denote an irreducible unitary representation of G. Then Schur's lemma implies that the action of Z on $\mathfrak{H}$ is by a unitary character; i.e. there exists a continuous character $\omega_\pi : Z \to \mathbf{S}^1$ such that $\pi(z)x = \omega_\pi(z)x$ for any $x \in \mathfrak{H}$: this is called the *central character* of π.

Remark: Let $(\mathfrak{H}, \langle \cdot, \cdot \rangle)$ be a Hilbert space. A convenient way to check that a unitary representation $(\rho, \mathfrak{H})$ is irreducible is to prove that any G-invariant continuous inner product $\langle \cdot, \cdot \rangle_2$ on $\mathfrak{H}$ is a multiple of $\langle \cdot, \cdot \rangle$. Indeed, if ρ contains a nonzero invariant closed proper subspace $\mathfrak{H}_0$ then under the decomposition $\mathfrak{H} = \mathfrak{H}_0 \oplus \mathfrak{H}_0^\perp$ we can change the inner product on $\mathfrak{H}_0^\perp$ by positive scalars while leaving the one on $\mathfrak{H}_0$ unchanged and this preserves the G-invariance property. But such a modification inner product on $\mathfrak{H}$ is clearly not a scalar multiple of the given one, so no such $\mathfrak{H}_0$ exists.

Conversely, if $(\rho, \mathfrak{H})$ is irreducible, and $\langle \cdot, \cdot \rangle_2$ is a G-invariant inner product on $\mathfrak{H}$, let $\mathfrak{H}'$ be the Hausdorff completion of $(\mathfrak{H}, \langle \cdot, \cdot \rangle_2)$: the natural embedding $(\mathfrak{H}, \langle \cdot, \cdot \rangle) \to (\mathfrak{H}', \langle \cdot, \cdot \rangle_2)$ is continuous, G-equivariant, with dense image. By Schur's lemma, it is a scalar multiple of an isometry: this proves our contention.

Lemma 3.1. *Let $(\rho, \mathfrak{H})$ be a unitary representation of G, and let $\mathfrak{H}'$ be a closed G-invariant subspace. If $(\rho, \mathfrak{H})$ is a Hilbert sum of irreducible representations, then so are $(\rho, \mathfrak{H}')$ and $(\rho, \mathfrak{H}/\mathfrak{H}')$.*

Proof: Using duality and/or orthogonal complements, it suffices to treat $\mathfrak{H}/\mathfrak{H}'$. Let's write:

$$\mathfrak{H} = \widehat{\bigoplus_{i \in \mathrm{I}}} \mathfrak{H}_i$$

where $\mathfrak{H}_i$ is an irreducible closed G-subspace of $\mathfrak{H}$ (the set of index I is at most countable if $\mathfrak{H}$ is separable, which will be the case in all our applications). We can also assume $\mathfrak{H}/\mathfrak{H}' \neq 0$.

The projection p onto $\mathfrak{H}/\mathfrak{H}'$ is G-equivariant, so $\mathfrak{H}/\mathfrak{H}'$ is spanned (in the Hilbert sense) by the $p(\mathfrak{H}_i)$ $(i \in I)$. In particular, some $p(\mathfrak{H}_i)$ is nonzero. But this projection is a closed G-invariant subspace of $\mathfrak{H}/\mathfrak{H}'$, so the set $\mathscr{X}$ of collections of pairwise orthogonal closed G-invariant irreducible subspaces of $\mathfrak{H}/\mathfrak{H}'$ is non-empty. By Zorn's Lemma there is a maximal element in $\mathscr{X}$, and the corresponding Hilbert direct sum is a closed G-invariant subspace W of $\mathfrak{H}/\mathfrak{H}'$. We just have to rule out the possibility that it is a proper subspace. If so, then clearly its orthogonal complement (in $\mathfrak{H}/\mathfrak{H}'$) contains no closed irreducible G-invariant subspace, so by replacing $\mathfrak{H}'$ with the preimage in $\mathfrak{H}$ corresponding to W we arrive at the case when the nonzero $\mathfrak{H}/\mathfrak{H}'$ contains no irreducible G-invariant closed subspaces. It has already been seen that such a situation cannot occur. qed

3.2. In some common situations, unitary representations are Hilbert sums of irreducibles representations: this is the content of the next theorems.

Theorem 3.2. *Let G be a compact group. Then any unitary representation is a Hilbert sum of irreducible representations. Furthermore any irreducible representation is finite dimensional.*

References: [Wal] prop. 1.4.1 and 1.4.2; [Ro] chapter 5 or the excellent [BR] chapter 7 for instance.

Remark: Let $(\rho, \mathfrak{H})$ be a unitary representation of G, and K be a compact subgroup. One can therefore write:

$$\mathfrak{H} = \widehat{\bigoplus_{i \in \mathrm{I}}} \mathfrak{H}_i$$

where each $\mathfrak{H}_i$ is a K-irreducible closed subspace of $\mathfrak{H}$. This decomposition is not unique (think of the trivial representation, for which any Hilbert basis

provides such a decomposition), and two $\mathfrak{H}_i$'s may be unitarily K-equivalent. One usually rewrites the decomposition as follows: for each (isomorphism class of) irreducible representation π of K , let I_π be the set of $i \in \mathrm{I}$ for which $(\rho_{|_K}, \mathfrak{H}_i)$ is equivalent to π. The cardinal number $m_\pi = \mathrm{card}(\mathrm{I}_\pi)$ is the *multiplicity* of π in $\rho_{|_K}$: by Schur lemma, this cardinal number is independent of the decomposition we started with. One writes $\mathfrak{H}(\pi) = \bigoplus_{i\in \mathrm{I}_\pi} \mathfrak{H}_i = m_\pi \rho$, and the above Hilbert sum is written:

$$\mathfrak{H} = \widehat{\bigoplus_{\pi\in\widehat{K}}} \mathfrak{H}(\pi) = \widehat{\bigoplus_{\pi\in\widehat{K}}} m_\pi \pi.$$

One says that ρ is K-*admissible* if m_π is a finite cardinal for each $\pi \in \widehat{K}$. We'll see later on that any irreducible unitary representation of a connected reductive group is admissible (for K a maximal compact subgroup in the archimedean case, and maximal compact open subgroup in the non-archimedean case).

3.3. The next examples require the use of the integration in topological vector spaces. A thorough treatment can be found in Bourbaki, Integration, chap VI, §1,2 and chap VII, §2 for the application on representations; [War] section 4.1.1; [Ro] section 6 for some comments. Let $(\pi, \mathfrak{H})$ be a unitary representation of a locally compact group G (so $\mathfrak{H}$ is a Hilbert space, though to integrate continuous vector-valued functions it suffices to assume that $\mathfrak{H}$ is locally convex and quasi-complete). Let $f \in \mathcal{C}_c(G)$, $v, w \in \mathfrak{H}$, one can consider the absolutely converging integral:

$$l_v(w) := \int_G f(g) \langle \pi(g)v, w\rangle \, dg$$

The mapping $w \mapsto l_v(w)$ is continuous and linear, therefore by Riesz' representation theorem it defines an element of $\mathfrak{H}$ denoted

$$\pi(f)v := \int_G f(g)\pi(g)v dg.$$

It is clearly linear in f and v, continuous as $\|\pi(f)v\| \leq \|f\|_1\|v\|$ and can be extended by density to $L^1(G)$ (actually even to the space of compactly supported complex measures, cf. Bourbaki): in particular, one checks easily that $f \in L^1(G) \mapsto \pi(f) \in \mathrm{End}(\mathfrak{H})$ is a continuous homomorphism of Banach algebras.

REMARK: It is sometimes convenient to consider a continuous function f whose support is contained in a compact subgroup K of G. If K is negligible in G, then $\pi(f)$ as defined above is zero. However, the same arguments shows that the integral $\int_K f(k)\pi(k)vdk$ is absolutely convergent: by an abuse of notations, we will denote this integral $\pi(f)v$. As soon as the Haar measures are suitably normalized, this defines the same operator in the case K is also open, so we hope this won't cause any confusion.

Theorem 3.3. *Let $(\pi, \mathfrak{H})$ be a unitary representation of G. Assume the existence of a delta-sequence $(f_n)_{n\in\mathbf{N}}$ in $\mathcal{C}_c(G)$, i.e. satisfying:*

$$\mathrm{supp}(f_{n+1}) \subset \mathrm{supp}(f_n),\ \bigcap_{n\geq 1} \mathrm{supp}(f_n) = \{e\}$$

$$\forall n \in \mathbf{N},\ \forall g \in G,\ f_n(g) = f_n(g^{-1}), f_n \geq 0, \int_G f_n = 1$$

such that the operator $\pi(f_n)$ is compact for all n. Then $(\pi, \mathfrak{H})$ is a Hilbert sum of irreducible representations, each occuring with finite multiplicities.

References: [Wal] proposition 1.4.1, [L] I §3. Note that the invariance of the f_n under $g \mapsto g^{-1}$ insures that $\pi(f_n)$ is self-adjoint: the proof uses the spectral decomposition of such operators.

Remark: This theorem is fundamental in the theory of automorphic forms: the most common proofs that the space of cusp forms splits as a sum of irreducible representations is based on it – though Jacquet-Langlands seem to have a purely algebraic proof of this fact.

Remark: Let G be a locally compact group. One says that G (actually its stellar algebra: see [Dix], 13.9) is *liminal* if for any $(\pi, \mathfrak{H})$ irreducible unitary representation of G, and any $f \in \mathcal{C}_c(G)$, $\pi(f)$ is compact. We'll see later that all reductive groups over locally compact fields are liminal, and to what extent this plays a role in the tensor product theorem.

Theorem 3.4. *Let G be a locally compact group, K a compact subgroup of G, and $(\pi, \mathfrak{H})$ a unitary representation of G. Assume that π is K-admissible. Then there exists a delta-sequence satisfying the condition of the previous theorem, and therefore $(\pi, \mathfrak{H})$ splits as a Hilbert direct sum of irreducible representations.*

Proof: (cf. [Bo2], 5.9 corollaire) If $\rho \in \widehat{K}$ occurs in π, denote its character χ_ρ: by hypothesis $\pi(\chi_\rho)$ is compact. As any central function f of K is a uniform limit of linear combinations of characters (cf. [Ro], 7.1, proposition), so $\pi(f)$ is compact as well (the subspace of compact operators is closed in $\mathrm{End}(\mathfrak{H})$ for the topology of uniform convergence on bounded sets). To conclude, one uses a delta-sequence made of central functions (by averaging over K of course), and one applies the previous theorem. qed

4. The case of reductive groups

In this section, let G be a reductive algebraic group over a local field F (say $F = \mathbf{R}$ or $\mathbf{Q}_p$ for some prime p). One denotes $\mathfrak{g}$ its Lie algebra. Let K be a compact subgroup of $G(F)$ such that:

- if F is archimedean, K is a maximal compact subgroup of $G(F)$ (e.g. $K = \mathbf{O}_2(\mathbf{R})$ if $G = \mathbf{GL}_2, F = \mathbf{R}$)
- if F is non-archimedean, K is open (e.g. $K = \mathbf{GL}_2(\mathbf{Z}_p)$ if $G = \mathbf{GL}_2, F = \mathbf{Q}_p$)

THEOREM 4.1. *Let $(\pi, \mathfrak{H})$ be an irreducible unitary representation of $G(F)$. Then $(\pi, \mathfrak{H})$ is K-admissible.*

References: for F archimedean, cf. [Wal] theorem 3.4.10, [Bo2] théorème 5.27. In the non-archimedean case, it is quoted by Cartier in [Cor] and it is discussed in the unpublished notes of Garrett [Ga1]. As we are mainly interested in the case of $G = \mathbf{GL}(2)$, refer to [Su] theorem 5.1. for the real case, and to [BH] in the p-adic case, where the smooth representations are completely classified – so that one is left to observe the admissibility.

COROLLARY 4.1.1. *Let $(\pi, \mathfrak{H})$ be an irreducible unitary representation of $G(F)$, and $f \in \mathcal{C}_c(G)$. Then $\pi(f)$ is compact.*

References: Théorème 5.27 in [Bo2] for the real case. In the p-adic case, $\mathcal{C}_c^\infty(G)$ is dense in $\mathcal{C}_c(G)$ (for its natural inductive limit topology, stronger than the uniform convergence): as the subspace of compact operators is closed in $\mathrm{End}(\mathfrak{H})$, it suffices to prove the claim for $f \in \mathcal{C}_c^\infty(G)$. But for such an f, it is immediate that one can find a compact open subgroup K_f of G such that $f(kgk^{-1}) = f(g)$ for any $g \in G$, $k \in K_f$, in which case one concludes as in the proof of theorem 3.4.

REMARK: This proves that the (stellar algebra of) reductive groups are liminal, as claimed above.

Before we state the next corollary, which will be useful in our discussion of the tensor product theorem, let's recall that given two Hilbert spaces $\mathfrak{H}_1, \mathfrak{H}_2$, the bilinear map induced by

$$\langle x_1 \otimes x_2, y_1 \otimes y_2 \rangle := \langle x_1, y_1 \rangle_1 \langle x_2, y_2 \rangle_2$$

provides $\mathfrak{H}_1 \otimes \mathfrak{H}_2$ with a non-degenerate inner product, whose completion is denoted $\mathfrak{H}_1 \widehat{\otimes} \mathfrak{H}_2$: cf [Bour-EVT], chap V, §3, No 1 and 2. Let G_1, G_2 be two locally compact groups, and let $(\pi_i, \mathfrak{H}_i)$ $(i = 1, 2)$ be two unitary representations. Then $\pi_1 \widehat{\otimes} \pi_2$ denotes the unitary representation of $G_1 \times G_2$ on $\mathfrak{H}_1 \widehat{\otimes} \mathfrak{H}_2$ deduced from the representation on the pre-Hilbert space $\mathfrak{H}_1 \otimes \mathfrak{H}_2$, itself induced by:

$$(\pi_1 \otimes \pi_2)(g_1, g_2)(x_1 \otimes x_2) = \pi_1(g_1)x_1 \otimes \pi_2(g_2)x_2.$$

Let's briefly justify this is a unitary representation: first, for any $g_1 \in G_1, g_2 \in G_2$, the operator $\pi_1(g_1) \otimes \pi_2(g_2)$ is unitary ([Bour-EVT], V, §4, No 1, proposition 3 and the paragraph following proposition 2); as for the continuity, given that $G = G_1 \times G_2$ acts by unitary operators, it suffices to prove that the mappings $g \in G \mapsto \pi(g)v$ are continuous for v in a total subset of $\mathfrak{H}_1 \widehat{\otimes} \mathfrak{H}_2$ (cf [War] section 4.1.1 page 219): if we take this total subset to be $\{x_1 \otimes x_2 \ : \ x_1 \in \mathfrak{H}_1, x_2 \in \mathfrak{H}_2\}$, our contention is clear.

It is easy to see that $\pi_1 \widehat{\otimes} \pi_2$ is irreducible if π_1, π_2 are. Indeed, let Q be a $G_1 \times G_2$-invariant continuous inner product on $\mathfrak{H}_1 \otimes \mathfrak{H}_2$. Fix two nonzero vectors $x_2, y_2 \in \mathfrak{H}_2$: then the inner product on $\mathfrak{H}_1$ defined by $(x_1, y_1) \mapsto Q(x_1 \otimes x_2, y_1 \otimes y_2)$ is continuous and G_1-invariant, so is equal to $\langle \cdot, \cdot \rangle_1$ up

to some constant by irreducibility of π_1. One determines the constant the same way, by varying x_2, y_2. Conversely:

Corollary 4.1.2. *Let G_1, G_2 be two reductive groups over two local fields (maybe distinct). Then any irreducible representation π of $G = G_1 \times G_2$ is equivalent to a representation of the shape $\pi_1 \widehat{\otimes} \pi_2$, with π_1, π_2 irreducible.*

Reference: [Dix] proposition 13.1.8, where it is proven that if at least one of G_1, G_2 are of type 1, then the conclusion holds (see also [GGP] appendix to chapter 2 and [Ro], section 20). It can be shown (cf. [Dix], theorem 5.5.2 and 13.9.4) that a group is of type 1 as soon as its stellar algebra is liminal, which is the case here, by the corollary 4.1.1.

Remark: The previous corollary was stated only in the case of reductive groups: it is of course true in the generality of type 1 groups, as the references justify it.

4.1. **Smooth vectors and $(\mathfrak{g}, K)$-modules.** References: [Bu] chapter 2, [Wal1], and [Wal] chapter 3. As we mentionned earlier, there are also more algebraic counterparts of representation theory. In the case of archimedean Lie groups, $(\mathfrak{g}, K)$-modules play an important role. Let G be an archimedean reductive Lie group, $\mathfrak{g}$ its complex Lie algebra, K a maximal compact subgroup.

One can attach canonically to $\mathfrak{g}$ an associative unitary algebra $\mathcal{U}(\mathfrak{g})$ called the (complexified) universal enveloping algebra, which gives rise to differential operators acting on $\mathcal{C}_c^\infty(G)$. We will denote $\mathfrak{z}$ the center of $\mathcal{U}(\mathfrak{g})$ (if G is of inner type, this is also the set of elements z in $\mathcal{U}(\mathfrak{g})$ such that $\mathrm{Ad}(g)z = z$ for any g in G, cf [Wal] 3.4.1: this is the case for $\mathbf{GL}_n(\mathbf{R})$), which is finitely generated, and generalizes the Laplace-Beltrami operator: cf [Wal] section 0.4.

A $(\mathfrak{g}, K)$-*module* is a complex vector space V (without topology), together with

- a structure of a K-module, continuous in the following sense: if $v \in V$, then there exists a finite dimensional subspace W_v such that $Kv \subset W_v$ and the mapping $K \to \mathrm{Aut}(W_v)$ is continuous (therefore analytic),
- a structure of a $\mathfrak{g}$-module,

such that:

(1) $k \cdot X \cdot v = (\mathrm{Ad}(k)X) \cdot k \cdot v$ for $k \in K, X \in \mathfrak{g}, v \in V$,
(2) $\frac{d}{dt}(\exp(tX)v)_{|t=0} = Xv$ for $v \in V$ and X in the Lie algebra $\mathfrak{k}$ of K.

In these conditions, one can prove that V is a semisimple K-module (cf. [Wal] lemma 3.3.3). The $(\mathfrak{g}, K)$-module V is *admissible* if the ρ-isotypic subspace $V(\rho)$ is finite-dimensional for any $\rho \in \widehat{K}$.

Fundamental example: Let $(\pi, \mathfrak{H})$ be a unitary representation of G. Let $\mathfrak{H}^\infty$ be the subspace of smooth vectors (i.e. the vectors $v \in \mathfrak{H}$ such that $g \in G \mapsto \pi(g)v$ is smooth). The real Lie algebra $\mathfrak{g}_\mathbf{R}$ acts on $\mathfrak{H}^\infty$ by $d\pi(X)v =$

$\frac{d}{dt}(\exp(tX)v)_{|t=0}$[†], hence an action of $\mathfrak{g}$. Gårding's theorem states that $\mathfrak{H}^\infty$ is dense in $\mathfrak{H}$ (exercise: use a δ-sequence of smooth functions to prove it, cf [Wal] section 1.6). Write $\mathfrak{H} = \widehat{\bigoplus}_{\rho\in\widehat{K}}\mathfrak{H}(\rho)$: one can prove that $\mathfrak{H}^\infty \cap \mathfrak{H}(\rho)$ is dense in $\mathfrak{H}(\rho)$ for any $\rho \in \widehat{K}$. Define $\mathfrak{H}_K := \bigoplus_{\rho\in\widehat{K}}(\mathfrak{H}(\rho) \cap \mathfrak{H}^\infty)$. Then $\mathfrak{H}_K$ is stable under the action of $K, \mathfrak{g}$, satisfies the aforementioned compatibilities and is called the $(\mathfrak{g}, K)$-module *associated* to the unitary representation $(\pi, \mathfrak{H})$. By construction, $\mathfrak{H}_K$ is dense in $\mathfrak{H}$. One can prove that $\mathfrak{H}_K$ is irreducible (=contains no algebraic submodule) if and only if the representation π is (topologically) irreducible, thanks to this density: cf. [Wal], theorem 3.4.11 – this uses the admissibility of π.

REMARK: Note that if $(\pi, \mathfrak{H})$ is admissible, as $\mathfrak{H}^\infty \cap \mathfrak{H}(\rho)$ is dense in $\mathfrak{H}(\rho)$, it must be *equal* to it. This implies that given an irreducible (necessarily admissible) unitary representation of a reductive group G, its associated $(\mathfrak{g}, K)$-module is actually $\bigoplus_{\rho\in\widehat{K}} \mathfrak{H}(\rho)$, and that the K-finite vectors of $\mathfrak{H}$ are smooth.

REMARK: A $(\mathfrak{g}, K)$-module does not afford a representation of G. However, one can define an "extension" of G, called the *Hecke algebra* and denoted $\mathcal{H}_G$, such that $(\mathfrak{g}, K)$-modules correspond naturally to $\mathcal{H}_G$-modules: see [Bu] proposition 3.4.4.

REMARK: There is a version of the Schur lemma for irreducible $(\mathfrak{g}, K)$-modules: cf [Wal] lemma 3.3.2. One can say a bit more in the case of an irreducible unitary representation of $G(\mathbf{R})$ for reductive G: the center of the universal algebra $\mathfrak{z}$ acts on $\mathfrak{H}^\infty$ by a character (here this means an homomorphism of $\mathbf{C}$-algebras $\chi : \mathfrak{z} \to \mathbf{C}$), this is the content of lemma 1.6.5 of [Wal].

REMARK: The complex conjugation on $\mathfrak{g}$ extends to an conjugate-linear anti-automorphism on $\mathcal{U}(\mathfrak{g})$ (cf [Wal] 1.6.5) denoted $x \mapsto x^\star$. The proof of lemma 1.6.5 (ibid.) implies that if $x \in \mathcal{U}(\mathfrak{g})$, then for any $v, w \in \mathfrak{H}$, with $(\pi, \mathfrak{H})$ unitary representation of G, $\langle d\pi(x)v, w\rangle = \langle v, d\pi(x^\star)w\rangle$. In particular if $x = x^\star$, then $d\pi(x)$ is self-adjoint. This applies to the Laplace-Beltrami operator Δ of $\mathbf{SL}_2(\mathbf{R})$ acting for example on $L^2(\mathbf{SL}_2(\mathbf{Z})\backslash\mathbf{SL}_2(\mathbf{R}))$, giving a representation-theoritic proof of such self-adjointness in this case, usually proved by Green's identity, cf [Bu] section 2.1.

REMARK: About K and $\mathfrak{z}$-finiteness, useful in the context of automorphic forms. If $(\pi, \mathfrak{H})$ is a unitary representation, a vector v is *K-finite* if $\pi(K)v$ is finite dimensional: this makes sense for any vector in the representation. If v is a smooth vector, then v is $\mathfrak{z}$-finite if $d\pi(\mathfrak{z})v$ is finite dimensional. However, it is technically important to define it for non-smooth vectors as well: this is

[†]The limit in consideration is with respect to the norm of $\mathfrak{H}$: when $\mathfrak{H}$ is a space of functions, the derivative can also taken with respect to the pointwise convergence, which may not be coherent with the latter. For instance, the smooth vectors in $L^2(\mathbf{R})$ is not $\mathcal{C}^\infty(\mathbf{R})$!

way distributions play an important role in the theory of automorphic forms, often implicitly. In this context, a vector $v \in \mathfrak{H}$ defines a (vector-valued) distribution $T_v : \mathcal{C}_c^\infty(G) \to \mathfrak{H}$ by:

$$T_v(\varphi) = \int_G \varphi(g)\pi(g)v dg.$$

One says that v is $\mathfrak{z}$-finite (as a distribution) if $\mathrm{span}\{zT_v : z \in \mathfrak{z}\}$ is a finite-dimensional subspace of $\mathfrak{H}$-valued distributions, where xT_v is the distribution defined for $x \in \mathcal{U}(\mathfrak{g})$ by:

$$(xT_v)(\varphi) := \int_G (\varphi * \check{x})(g)\pi(g)dg.$$

(Here and below, for $x \in \mathcal{U}(\mathfrak{g})$ the notation $\varphi * \check{x}$ denotes the action of x on $\mathcal{C}_c^\infty(G)$ arising from the action of $\mathfrak{g}$ via differential operators.) Note that if v is smooth then $xT_v = T_{d\pi(x)v}$ and that in the case where $(\pi, \mathfrak{H})$ is the right regular representation of $L^2(G)$, $f \in L^2(G)$ is $\mathfrak{z}$-finite in the above sense if and only if the real valued representations $\varphi \mapsto \int_G (\varphi * \check{z})gf(g)dg$ span, when z varies in $\mathfrak{z}$, a finite dimensional subspace of (real valued) distributions (by using the right regular representation on $\mathcal{C}_c^\infty(G)$).

4.2. **Smooth representations of non-archimedean groups.** References [BH] chapter 1, [Bu] chapter 4 for a thorough discussion of this topic. This is the p-adic counterpart of the preceeding paragraph. Let G be a totally disconnected locally compact group, K an open compact subgroup. A *smooth representation* of G is a vector space V together with an group homomorphism $\pi : G \to \mathrm{Aut}(V)$ such that

(1) any $v \in V$ is smooth, i.e. the subgroup $\{g \in G \ : \ \pi(g)v = v\}$ is compact and open in G.

In this situation, the restriction of the representation π to K is semisimple (cf. [BH] lemma 2.2). It is said to be *admissible* if furthermore the space of K-fixed vectors V^K is finite dimensional: this implies that one can write

$$V = \bigoplus_{\rho \in \widehat{K}} V(\rho)$$

where each $V(\rho)$ is finite-dimensional.

Fundamental example: Let $(\pi, \mathfrak{H})$ be a unitary representation of G; denote by $\mathfrak{H}^\infty$ the subspace of smooth vectors in $\mathfrak{H}$, which is stable under G. Then the corestriction of π to $\mathfrak{H}^\infty$ is a smooth representation of G.

Note that $\mathfrak{H}^\infty$ is dense in $\mathfrak{H}$: indeed, if $v \in V$, then $\pi(f)v \in \mathfrak{H}^\infty$ for any $f \in \mathcal{C}_c^\infty(G)$. Let $\mathfrak{C}$ be the filter generated by open and compact neighbourhoods of the identity and let $f_{\mathfrak{c}}$ be the characteristic function of $\mathfrak{c} \in \mathfrak{C}$: then $\pi(f_{\mathfrak{c}})v \to_{\mathfrak{C}} v$, as claimed. This density implies that given an admissible unitary representation $(\pi, \mathfrak{H})$, then $(\pi, \mathfrak{H})$ is irreducible if and only if $(\pi, \mathfrak{H}^\infty)$ is algebraically irreducible.

REMARK: Smooth representations of G are in one-to-one correspondence with smooth representations of the Hecke algebra of G: cf [BH] section 1.

CASE OF $\mathbf{GL}_2$: SPHERICAL REPRESENTATIONS. Here F denotes a non-archimedean field, and $\mathcal{O}_F$ its integers with maximal ideal $\mathfrak{p}_F$, ϖ a uniformizer and q_F the cardinality of the residue field. As we mentioned earlier, the smooth irreducible representations of $\mathbf{GL}_2(F)$ are classified (up to equivalence), and fall into three families: principal series, special representations and supercuspidals (see [BH]) . We will need later a few facts on unramified representations (as defined in the following result):

THEOREM 4.2. *Let (π, V) be a smooth irreducible representation of $\mathbf{GL}_2(F)$ which is unramified in the sense that it contains nonzero spherical vectors; i.e.,*

$$V^{\mathbf{GL}_2(\mathcal{O}_F)} := \{v \in V \ : \ \pi(k)v = v \text{ for all } k \in \mathbf{GL}_2(\mathcal{O}_F)\} \neq \{0\}$$

Then (π, V) is equivalent to an unramified principal series representation, and furthermore the space $V^{\mathbf{GL}_2(\mathcal{O}_F)}$ is one-dimensional.

This means that $\pi \cong \pi(\chi_1, \chi_2)$ for some unramified quasi-characters of $F^\times$. The one-dimensionality result comes from the fact that the spherical Hecke algebra is commutative (cf [Bu] theorem 4.6.2)

A natural question is, given a unitary irreducible representation π of $\mathbf{GL}_2(F)$, how to determine the characters χ_1, χ_2 from π: this leads to the introduction of Hecke operators in this local setting.

First of all, an unramified quasi-character χ of F can be written $\chi(x) = |x|^t$, for some $t \in \mathbf{C}$ (uniquely determined modulo $2i\pi \log(q_F)^{-1}\mathbf{Z}$). As it is known that $\pi \cong \pi(\chi_1, \chi_2)$ and $\pi \cong \pi(\chi_2, \chi_1)$ are unitarily equivalent, it suffices to determine the set $\{q_F^{t_1}, q_F^{t_2}\}$ of complex numbers (here the t_i's are actually imaginary, as the quasi-characters χ_1, χ_2 are unitary). By looking at the central characters, one has for any $x \in F^\times$:

$$\omega_\pi(x) = \chi_1(x)\chi_2(x)$$

so this gives a condition on $q_F^{t_1} q_F^{t_2}$.

To determine completely our set, it suffices to get a condition on the sum $q_F^{t_1} + q_F^{t_2}$. Let φ_0 denote the characteristic function of the (compact and open) subset $\mathbf{GL}_2(\mathcal{O}_F)\left[\begin{smallmatrix} \varpi & 0 \\ 0 & 1 \end{smallmatrix}\right]\mathbf{GL}_2(\mathcal{O}_F)$; then for any vector v, $\pi(\varphi_0)v$ is spherical (if non-zero). Thus, in our setting, if one denotes v_0 a non-zero spherical vector in $\pi(\chi_1, \chi_2)$, $\pi(\varphi_0)v$ and v are colinear, more precisely – and this will end our discussion:

$$\pi(\varphi_0)v = q_F^{1/2}(q_F^{t_1} + q_F^{t_2})v.$$

For a proof, see for example [Bu] proposition 4.6.6: the standard notation for the operator $\pi(\varphi_0)$ is $T(\mathfrak{p}_F)$ or $T_{\mathfrak{p}_F}$ – we'll see later that the Hecke operators introduced in section 2.3 "correspond" to these $T(\mathfrak{p}_F)$'s once the adeles are introduced.

5. The adelization of a modular form and adelic Hecke operators

In this section, we use the same notation as in section 2.

If G is an algebraic group over $\mathbf{Q}$ (we'll use $G = \mathbf{GL}_2$ only), we'll denote $g = (g_v)_{v \leq \infty}$ an element of $G(\mathbb{A}_{\mathbf{Q}})$. Given a place v of $\mathbf{Q}$, and an element $g_v \in G(\mathbf{Q}_v)$, we'll denote also g_v the element of $G(\mathbb{A}_{\mathbf{Q}})$ whose component at v is g_v, and at $w \neq v$ is 1. We'll denote sometimes g_f the "finite" part of g (i.e. $(g_f)_\infty = 1$ and $(g_f)_p = g_p$ for p prime). For $G = \mathbf{GL}_2$, we recall that a maximal compact subgroup of $G(\mathbb{A}_{\mathbf{Q}})$ is $K = \prod_{v \leq \infty} K_v$, with $K_\infty = \mathbf{O}_2(\mathbf{R})$ and $K_p = \mathbf{GL}_2(\mathbf{Z}_p)$ for p prime; the center of $G(\mathbb{A}_{\mathbf{Q}})$ will always be denoted $Z(\mathbb{A}_{\mathbf{Q}})$. If $N \in \mathbf{N}$, we'll denote $K_0(N)$ the subgroup of K_f made of matrices whose lower left entry is in $N\widehat{\mathbf{Z}}$ (so the component at infinity is 1).

5.1. From a classical modular form to an automorphic form on $\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$.

References: [KL] section 12.2, [G] §3, [Bu] section 3.6, [BCSGKK] section 7.

Let f be a modular form of weight k, nebentypus χ and level N. The Dirichlet character χ is associated with a finite order idele class character of $\mathbb{A}_{\mathbf{Q}}^\times/\mathbf{Q}^\times$ denoted ω_χ called the adelization of χ: see [KL] section 12.1 for its construction. The strong apprimation theorem states that:

$$\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}}) = \mathbf{GL}_2(\mathbf{Q})\mathbf{GL}_2^+(\mathbf{R})K_0(N).$$

This means that any element g in $\mathbf{GL}_2(\mathbb{A})$ can be (non-uniquely) written $g = \gamma h_\infty k$, with $\gamma \in \mathbf{GL}_2(\mathbf{Q}), h_\infty \in \mathbf{GL}_2^+(\mathbf{R}), k \in K_0(N)$ (in other words the continuous map $\mathbf{GL}_2(\mathbf{Q}) \times \mathbf{GL}_2^+(\mathbf{R}) \times K_0(N) \to \mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$ is surjective). See [KL] section 6.3 for an elementary proof in this setting.

One can prove as a consequence that $\mathrm{vol}(Z(\mathbb{A}_{\mathbf{Q}})\mathbf{GL}_2(\mathbf{Q})\backslash\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})) < \infty$: cf [KL] section 7.11. We will still denote ω_χ the character on $K_0(N)$ defined by the evaluation of ω_χ at the lower right entry.

Definition 5.1. *Let f be a modular form of weight k, nebentypus χ and level N. The* adelization *of f is the function $\varphi_f : \mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}}) \to \mathbf{C}$ defined by:*

$$\varphi_f(g) = j(h_\infty, i)^{-k} f(h_\infty \cdot i)\omega_\chi(k)$$

where:

1. *$h_\infty \in \mathbf{GL}_2^+(\mathbf{R}), k \in K_0(N)$ are chosen so that $g = \gamma h_\infty k$ for some $\gamma \in \mathbf{GL}_2(\mathbf{Q})$,*
2. *for any $z \in \mathbf{C} - \mathbf{R}$, $j(h_\infty, z) = \det(h_\infty)^{-1/2}(cz + d)$, if one write $h_\infty = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$*

Remarks:

- This is well defined (i.e. the number $\varphi_f(g)$ does not depend on the choices of γ, h_∞, k) because of the modularity of f: see section 12.2 of [KL].
- The function φ_f is continuous. Indeed, its restriction to the (open) subset $\gamma \mathbf{GL}_2^+(\mathbf{R}) K_0(N)$ (for any $\gamma \in \mathbf{GL}_2(\mathbf{Q})$) is continuous by definition, so by the "gluing lemma" φ_f is continuous on $\mathbf{GL}_2(\mathbb{A}_\mathbf{Q})$.
- For any $\gamma \in \mathbf{GL}_2(\mathbf{Q}), g \in \mathbf{GL}_2(\mathbb{A}_\mathbf{Q})$, $\varphi_f(\gamma g) = \varphi_f(g)$.
- For a fixed finite adelic point g_f, $g_\infty \mapsto \varphi_f(g_\infty, g_f)$ is smooth.
- For a fixed $g_\infty \in \mathbf{GL}_2^+(\mathbf{R})$, $g_f \mapsto \varphi_f(g_\infty, g_f)$ is locally constant on the finite adeles.

The last three points are obvious.

As the title of this section indicates, the function φ_f is actually an automorphic form on $\mathbf{GL}_2(\mathbb{A}_\mathbf{Q})$: we list below the properties statisfied by φ_f to inherit such a name: the proofs are to be found in [KL] or [G].

(1) ($\mathbf{GL}_2(\mathbf{Q})$-left invariance) For any $\gamma \in \mathbf{GL}_2(\mathbf{Q}), g \in \mathbf{GL}_2(\mathbb{A}_\mathbf{Q})$, one has: $\varphi_f(\gamma g) = \varphi_f(g)$.

(2) (K-finiteness) For $k_\infty = \left[\begin{smallmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{smallmatrix}\right] \in \mathbf{SO}_2(\mathbf{R}), k_f \in K_0(N), g \in \mathbf{GL}_2(\mathbb{A}_\mathbf{Q})$, $\varphi_f(g k_\infty k_f) = \omega(k_f)\exp(2\pi i k\theta)\varphi_f(g)$, where $k \in \mathbf{Z}$ is the weight of f. In the adelic setting, the condition of *K-finiteness* on φ_f means that the subspace $\mathrm{span}(R(g)\varphi_f \ : \ g \in K)$ is finite-dimensional. The link with the classical setting is that all finite-dimensional continuous representations of the circle group $\mathbf{SO}_2(\mathbf{R}) = \mathbf{R}/(2\pi\mathbf{Z})$ are direct sums of 1-dimensional representations with the character $\theta \mapsto \exp(ik\theta)$ for various $k \in \mathbf{Z}$.

(3) ($\mathfrak{z}$-finiteness) One has the differential equation: $\Delta\varphi_f = \frac{k}{2}\left(1 - \frac{k}{2}\right)\varphi_f$ (where the Casimir operator Δ acts on the infinite component). This implies that φ is Δ-finite: this, and the next item, implies that φ is $\mathfrak{z}$-finite, as the center of the universal algebra is generated by Δ and I). In other words, the subspace $\mathrm{span}(\varphi_f * \check{z} \ : \ z \in \mathfrak{z})$ is finite dimensional.

(4) (Action of the center) For any $z \in Z(\mathbb{A}_\mathbf{Q}), g \in \mathbf{GL}_2(\mathbb{A}_\mathbf{Q})$, $\varphi_f(zg) = \omega_\chi(z)\varphi_f(g)$.

(5) (Growth condition) For any norm $\|\cdot\|$ on $\mathbf{GL}_2(\mathbb{A}_\mathbf{Q})$, there exists a real number $A > 0$ such that: $\varphi_f(g) \ll \|g\|^A$. In other words, φ_f is moderate growth. This point is not obvious: see [Bo1] section 5, Borel-Jacquet in [Cor] and [Wal] for norms on Lie groups. It is simpler to prove that if f is a cusp form, then φ_f is actually bounded: this is because of the basic fact that f is cuspidal if and only if the mapping $g_\infty \in \mathbf{GL}_2^+(\mathbf{R}) \mapsto j(g_\infty, i)^{-k} f(g_\infty \cdot i)$ is bounded: see [KL] proposition 12.2.

(6) (Cuspidality) If f is a cusp form, then φ_f is cuspidal, in the sense that for any $g \in \mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$:

$$\int_{\mathbf{Q}\backslash\mathbb{A}_{\mathbf{Q}}} \varphi_f\left(\left[\begin{smallmatrix}1 & x\\ 0 & 1\end{smallmatrix}\right] g\right) dx = 0.$$

See [KL] proposition 12.3 for a proof (if the form is unramified, then [G] proves it as well, but for general levels, one has to use all the cusps).

We tried to list the properties so that they are easily modified to define an autmorphic form on a general reductive group G over $\mathbf{Q}$ (even over a number field); the cuspidality condition is more difficult to handle, as one has to write the vanishing condition on the unipotent radical of any parabolic $\mathbf{Q}$-subgroup of G.

It is important to notice that since φ_f is bounded for f cuspidal, $|\varphi_f|$ is square integrable on $Z(\mathbb{A})\mathbf{GL}_2(\mathbf{Q})\backslash\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$.

Remark: By using the strong approximation theorem, one can characterize the image of $\mathcal{S}_k(N,\chi)$ under this construction (which is clearly linear in f): refer to [KL] section 12.4.

5.2. **From a classical cuspidal modular form to a unitary automorphic representations of $\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$.** We keep the same notations, and refer to [G] for more details that we won't cover (chapter 5 is especially relevant).

Definition 5.2. *Let f be a cuspidal modular form of weight k, nebentypus χ and level N. The unitary automorphic representation attached to f is the restriction of the right regular representation of $\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$ on the closed subspace $\mathfrak{H}_f$ of $L^2_0(Z(\mathbb{A}_{\mathbf{Q}})\mathbf{GL}_2(\mathbf{Q})\backslash\mathbf{GL}_2(\mathbb{A}), \omega_\chi)$ defined by:*

$$\mathfrak{H}_f := \overline{\mathrm{span}(R(g)\varphi_f \ : \ g \in \mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}}))}$$

This unitary representation is denoted π_f.

Remark: see the appendix for a definition of the space of cuspidal functions $L^2_0(Z(\mathbb{A}_{\mathbf{Q}})\mathbf{GL}_2(\mathbf{Q})\backslash\mathbf{GL}_2(\mathbb{A}), \omega_\chi)$.

Remark: We chose to work with unitary automorphic representations; if instead one wishes to work with the more algebraic theory of (admissible) automorphic representations, one can attach to a modular form f the $\mathcal{H}_{\mathbf{GL}_2}$-submodule $\mathcal{H}_{\mathbf{GL}_2}\varphi_f$ of the space of automorphic forms on $\mathbf{GL}_2$ with central character ω_χ: here $\mathcal{H}_{\mathbf{GL}_2}$ denotes the adelic Hecke algebra, which is a restricted tensor product of the local Hecke algebras – cf [Bu] section 3.4.

The main result is the following:

Theorem 5.1. *Let f be a cuspidal modular form of weight k, nebentypus χ and level N. Assume that there exists a finite set of primes S such that f is a Hecke eigenform for the T_p, $p \notin S$. Then the unitary representation π_f is irreducible.*

We will sketch the proof below – references: [Bu] section 3.6, and [G] section 5.B. We will need the tensor product theorem and multiplicty one to achieve that.

5.3. **Hecke operators.** Let f be a cuspidal modular form of weight k, nebentypus χ and level N, and φ_f its adelization. We denote ω the adelization of the Dirichlet character χ. Let p be a prime not dividing q (for simplicity).

Let H_p be the compact open subset of $\mathbf{GL}_2(\mathbf{Q}_p)$ defined by

$$H_p = \mathbf{GL}_2(\mathbf{Z}_p)\left[\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right]\mathbf{GL}_2(\mathbf{Z}_p)$$

For $\varphi \in L^2_0(Z(\mathbb{A})\mathbf{GL}_2(\mathbf{Q})\backslash\mathbf{GL}_2(\mathbb{A}_\mathbf{Q}), \omega)$, we define:

$$\mathbb{T}_p(\varphi) = \int_{H_p} f(gk_p)dk_p.$$

By using the disjoint union decomposition:

$$H_p = \bigcup_{b=0}^{p-1}\left[\begin{smallmatrix} p & b \\ 0 & 1 \end{smallmatrix}\right]\mathbf{GL}_2(\mathbf{Z}_p) \cup \left[\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\right]\mathbf{GL}_2(\mathbf{Z}_p)$$

one proves easily that (cf [G] lemma 3.7):

$$\mathbb{T}_p(\varphi_f) = \varphi_{p^{1-k/2}T_p(f)}.$$

On the other hand, by using the disjoint union decomposition:

$$H_p = \bigcup_{b=0}^{p-1}\mathbf{GL}_2(\mathbf{Z}_p)\left[\begin{smallmatrix} 1 & b \\ 0 & p \end{smallmatrix}\right] \cup \mathbf{GL}_2(\mathbf{Z}_p)\left[\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right]$$

one sees that if φ is $\mathbf{GL}_2(\mathbf{Z}_p)$-invariant on the right then

$$\mathbb{T}_p(\varphi) = (p+1)\int_{\mathbf{GL}_2(\mathbf{Z}_p)} \varphi\left(gk_p\left[\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right]\right)dk_p.$$

The last integral can be modified in order to adelize the ramified Hecke operators. Reference: [M2] and [We] chapter VI.

REMARK: In [KL] section 13, it is explained how to use f to construct a smooth function ψ on $\mathbf{GL}_2(\mathbb{A}_\mathbf{Q})$ as a product $\psi = \prod_v \psi_v$, where ψ_∞ is integrable modulo the center for weights ≥ 3 and the finite components are smooth and compactly supported modulo the center: this is technically important in order to use the (relative) trace formula.

6. THE TENSOR PRODUCT THEOREM

In this section, we collect some facts leading to the statement and proof of the tensor product theorem. Again, our choice is to deal with unitary representations, for two reasons:

- Automorphic representations are not representations of $\mathbf{GL}_2(\mathbb{A}_\mathbf{Q})$, but of the Hecke algebra, which is difficult to define (see [Bu] section 3.4). It is simpler to define unitary representations in this context.

- Sooner or later in the theory, one really *needs* properties of Hilbert spaces, compact operators, trace class operators. The author is not sure to what extend a completely algebraic theory can do the job.

6.1. **A construction.** Let G denote the algebraic group $\mathbf{GL}_2$; all the notations introduced in the preceding section remain in force. For each place $v \leq \infty$ of $\mathbf{Q}$, let $(\pi_v, \mathfrak{H}_v)$ be a unitary representation of $G(\mathbf{Q}_v)$. Denote $\langle \cdot, \cdot \rangle_v$ the inner product of $\mathfrak{H}_v$. Let's assume that there exists a finite set of finite primes S_0 containing ∞ such that for any $v \notin S_0$, the space of K_v-fixed vectors is one-dimensional. For each place $v \notin S_0$, we choose a *unitary* vector in $\mathfrak{H}_v^{K_v}$, which we will denote ξ_v^0. We will construct a unitary representation π of $\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$ which is usually denoted

$$\pi = \widehat{\bigotimes_{v \leq \infty}} \pi_v$$

but one has to keep in mind that it might a priori depend on the choice of $\boldsymbol{\xi}^0 = (\xi_v^0)_{v \notin S_0}$ (and so $\widehat{\bigotimes_{v \leq \infty}}^{\xi} \pi_v$ would be a better notation).

Step 1: construction of the Hilbert space on which $\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$ will act: reference [Gui].

For each finite set S of primes containing S_0, one denotes by $\mathfrak{H}_S$ the prehilbert space

$$\mathfrak{H}_S = \bigotimes_{v \in S} \mathfrak{H}_v$$

By general properties of the tensor product of modules, one does not have to choose any order on the set of places. For two such sets S, T with $S \subset T$, there is a unique mapping $j_{S,T} : \mathfrak{H}_S \to \mathfrak{H}_T$ defined for each family $(x_v \in \prod_{v \in S} \mathfrak{H}_v)$ by:

$$j_{S,T}(\bigotimes_{v \in S} x_v) = \bigotimes_{v \in S} x_v \otimes \bigotimes_{v \in T - S} \xi_v^0.$$

It is obvious that these mappings are injective Put on $\mathfrak{H}_S$ the (positive definite) inner product $\langle \cdot, \cdot \rangle_S$ induced by

$$\langle \bigotimes_{v \in S} x_v, \bigotimes_{v \in S} y_v \rangle_S := \prod_{v \in S} \langle x_v, y_v \rangle_v$$

See section 4 for some facts on these tensor products. One sees immediately that the embeddings $j_{S,T}$ are isometric for these inner products. Denote $\mathfrak{H}^{\mathrm{alg}}$ the inductive limit of the system $(\mathfrak{H}_S, j_{S,T})$ (the directed set is the set of finite sets of primes, ordered by inclusion), and $j_S : \mathfrak{H}_S \to \mathfrak{H}^{\mathrm{alg}}$ the canonical embedding.

For $x, y \in \mathfrak{H}^{\mathrm{alg}}$, there exist a finite set of places S, and elements x_S, y_S of $\mathfrak{H}_S$ such that $x = j_S(x_S), y = j_S(y_S)$ and we define an inner product on $\mathfrak{H}^{\mathrm{alg}}$ by:

$$\langle x, y \rangle = \langle x_S, y_S \rangle_S$$

This is well-defined, and this makes each j_S an isometry. Finally, denote by $\mathfrak{H}$ the completion of $\mathfrak{H}^{\mathrm{alg}}$ for this inner product: this space is denoted in [Gui] as $\widehat{\bigotimes_{v\leq\infty}}^{\xi}\mathfrak{H}_v$.

Remark: In [Gui], it is proved that *canonically* there is no dependence on the vectors $\boldsymbol{\xi}^0$: this is Proposition 1.3 loc. cit., which can be applied as the space of K_v-fixed vectors is one-dimensional (for another choice $\boldsymbol{\xi}^1$, one has $\xi_v^1=\alpha_v\xi_v^0$ for a unique complex α_v of modulus one, so the hypothesis is trivially satisfied). Of course, the choice of the finite set S_0 is unimportant, by general properties of inductive limits.

Remark: If $x=\bigotimes_{v\in S}x_v$ is a vector in $\mathfrak{H}_S$, one often denotes its image in $\mathfrak{H}$ using the notation $\bigotimes_{v\in S}x_v\otimes\bigotimes_{v\notin S}\xi_v^0$: this is an abuse of language, as the latter makes sense only in the algebraic infinite tensor product $\bigotimes_{v\leq\infty}\mathfrak{H}_v$ as defined in Bourbaki, Algèbre, chap II, §3, No 9, but this is common.

Step 2: construction of the representation.

For each finite set S of primes containing S_0, $\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q},S})$ acts on $\mathfrak{H}_S$ via the unitary representation $\bigotimes_{v\in S}\pi_v$: unitarity and continuity of this representation has been checked in section 4. One sees at once that one gets an inductive system of unitary representations (of course by using the K_v-invariance of the ξ_v^0's), and so one gets an algebraic representation π^{alg} of $\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$ on $\mathfrak{H}^{\mathrm{alg}}$ by unitary operators. One can extend by uniform continuity each operator $\pi^{\mathrm{alg}}(g)$, for $g\in\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$, to a unitary operator on the completion $\mathfrak{H}$, which we denote π: π affords an algebraic representation of $\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$ by unitary operators, so we need only justify the continuity of this action.

It suffices to prove that the mappings $g\in\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})\mapsto\pi(g)x\in\mathfrak{H}$ are continuous for each x in a total subset of $\mathfrak{H}$ (see section 4), so it is sufficient to check this continuity for x of the shape $x=j_S(x_S)$ for some finite set S of primes containing S_0. Because of the topology on the adeles, it suffices to prove the continuity of $g\in\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q},T})\mapsto\pi(g)x\in\mathfrak{H}$ for T a fixed finite set of primes with $S\subset T$, but in this case one has:

$$\pi(g)x=j_T\left(\bigotimes_{v\in S}\pi_v(g_v)x_v\otimes\bigotimes_{v\in T-S}\pi_v(g_v)\xi_v^0\right)$$

and the continuity is clear.

We have therefore constructed from the data the unitary tensor product representation π, denoted in the litterature

$$\pi=\widehat{\bigotimes_{v\leq\infty}}\pi_v.$$

A slightly better notation would be $\widehat{\bigotimes_{v\leq\infty}}^{\boldsymbol{\xi}^0}\pi_v$ a priori, but because of the remark we made above, if one changes the family $\boldsymbol{\xi}^0$, one *canonically* gets a unitarily equivalent representation.

Remark: If furthermore all the local representations π_v are irreducible, then so is their unitary tensor product: this is a simple adaptation of an argument given in section 4. See [G], §4.C.

Remark: Let π be a unitary representation of $\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$ constructed as above from irreducible unitary representations π_v of $\mathbf{GL}_2(\mathbf{Q}_v)$. Then the restriction of π to $\mathbf{GL}_2(\mathbf{Q}_v)$ splits as a Hilbert direct sum of irreducible representations, all equivalent to π_v: this means that (the equivalence class of) π_v is uniquely determined by π.

Remark: If an irreducible unitary representation π of $\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$ is equivalent to a unitary tensor product as above, then π is admissible. Indeed, any $\rho\in\widehat{K}$ is equivalent to a unitary tensor product representation $\widehat{\bigotimes_v}\rho_v$, where ρ_v is an irreducible representation of K_v for each place v, almost all of which are trivial of dimension 1 ([Bu], lemma 3.3.1): each of these local representations appear with finite multiplicities in their respective space, which proves the claim.

Remark: If one is interested in the algebraic theory of the automorphic representations, one has to modify slightly the above construction to a more algebraic one: this is explained in Bump (ibid.).

6.2. **The tensor product theorem.** In this section, one is interested in a converse statement of the previous construction. We take $G=\mathbf{GL}_2$, but this would work mutatis mutandis for a reductive group over a number field, as these are liminal.

Theorem 6.1 (The tensor product theorem). *Let π be an irreducible unitary representation of $G(\mathbb{A}_{\mathbf{Q}})$. Then there exist a finite set S_0 of primes containing ∞, an irreducible unitary representation π_v of $G(\mathbf{Q}_v)$ for each place v such that π_v is spherical for $v\notin S_0$, and a unitary K_v-fixed vector ξ_v^0 for each $v\notin S_0$, so that π is equivalent to $\widehat{\bigotimes_{v\leq\infty}}^{\boldsymbol{\xi}^0}\pi_v$ for $\boldsymbol{\xi}^0=(\xi_v^0)_{v\notin S_0}$.*

Remark: see [Bu] section 3.4 for a statement and proof of the algebraic counterpart, as well as Cogdell in [CKM] lecture 3 for a statement without proof of the various versions of the tensor product theorem.

References: Depending on the strength of the statement, there are more or less difficult proofs of this result.

- Godement in [Go] §3.2 assumes furthermore that π is admissible. Under this assumption, he considers the restriction of π to $G(\mathbf{Q}_v)$, which is also admissible, and therefore splits as a Hilbert direct sum

of irreducible representations by theorem 3.4. As π is irreducible, the Schur lemma insures that any continuous operator in the space of π commuting with $\pi_{|_{G(\mathbf{Q}_v)}}$ and with the operators commuting with $\pi_{|_{G(\mathbf{Q}_v)}}$ are scalars, so that $\pi_{|_{G(\mathbf{Q}_v)}}$ is a factor representation.

As G is of type 1, this implies that $\pi_{|_{G(\mathbf{Q}_v)}}$ is isotypical, i.e. isomorphic to a Hilbert direct sum of equivalent representations (cf [Dix] or [Ro] section 20). Therefore, one has at one's disposal a family of irreducible unitary representations π_v of $G(\mathbf{Q}_v)$ for each place v, and the rest of the proof is a tedious construction allowing to "glue" together the local pieces. Note that along the way one chooses unitary K_v-fixed vectors, getting for each choice a factorization into a unitary tensor product – hence another justification in this context of the "independence" in the choice of $\boldsymbol{\xi}^0$.

- In [GGP] chapter 3 §3.3, there is a proof which does not make use of any admissibility condition. As a consequence, this proves that any irreducible representation of $G(\mathbb{A}_{\mathbf{Q}})$ is admissible, as explained in the previous subsection. Without a doubt, one could adapt Godement's arguments in order not to assume that π is admissible, as this is used only to find the local pieces π_v: [GGP] get these another way, yet the rest of the proofs are pretty close.
- If one is only interested in unitary cuspidal representations, one can prove first the admissibility of these, and use Godement's argument, or even the algebraic tensor product theorem on the space of K-finite vectors: in the last case, one gets a factorization into a restricted tensor product of smooth representations, which are unitarizable because π is.

  To prove the admissibility of an irreducible unitary cuspidal automorphic representation (π, V_π), that is (a unitary representation equivalent to a) $G(\mathbb{A}_{\mathbf{Q}})$-invariant irreducible closed subspace of the space $L^2_0(Z(\mathbb{A}_{\mathbf{Q}})\mathbf{GL}_2(\mathbf{Q})\backslash\mathbf{GL}_2(\mathbb{A}), \omega)$ for some unitary character ω of the idele class group, one can proceed as follows: let $\rho = \widehat{\bigotimes_v}\rho_v \in \widehat{K}$, where ρ_v is an irreducible representation of K_v for each place v. As K_f is totally disconnected, and ρ is finite dimensional, there exists an open compact normal subgroup K_1 of K_f such that the restriction of ρ to K_1 is trivial.

  One wants to prove that $V_\pi(\rho)$ is finite dimensional. The latter is contained in the space of K_1-fixed vectors in V_π which we denote $V_\pi^{K_1}$. Note right now that $V_\pi^{K_1}$ is stable under the restriction of the right regular representation of $\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$ to $\mathbf{GL}_2(\mathbf{R})$, so that it suffices to prove that $V_\pi^{K_1}(\rho_\infty)$ is finite dimensional.

  Consider $\varphi = \varphi_\infty\varphi_f \in \mathcal{C}_c^\infty(G(\mathbb{A}_{\mathbf{Q}}))$, where:
  - the function (defined on $\mathbf{GL}_2(\mathbb{A}_f)$) φ_f is the characteristic function of K_1,

– the archimedean component $\varphi_\infty \in \mathcal{C}_c^\infty(\mathbf{GL}_2(\mathbf{R}))$ is arbitrary.
The mapping $R(\varphi)$ defined for $f \in L_0^2(Z(\mathbb{A})\mathbf{GL}_2(\mathbf{Q})\backslash\mathbf{GL}_2(\mathbb{A}_\mathbf{Q}), \omega)$ by:

$$R(\varphi)(f)(x) = \int_{\mathbf{GL}_2(\mathbb{A}_\mathbf{Q})} \varphi(y) f(xy) dy \text{ for any } x \in \mathbf{GL}_2(\mathbb{A}_\mathbf{Q})$$

is a compact operator on $L_0^2(Z(\mathbb{A})\mathbf{GL}_2(\mathbf{Q})\backslash\mathbf{GL}_2(\mathbb{A}_\mathbf{Q}), \omega)$, and so is its restriction to V_π. But for $f \in V_\pi^{K_1}, x \in \mathbf{GL}_2(\mathbb{A}_\mathbf{Q})$, one has:

$$(R(\varphi)f)(x) = \mathrm{vol}(K_1) \int_{\mathbf{GL}_2(\mathbf{R})} \varphi_\infty(y_\infty) f(xy_\infty) dy_\infty$$

which is also a compact operator. As a consequence, theorem 3.3 implies that $V_\pi^{K_1}$ splits as a Hilbert direct sum of irreducible representations $(V_i)_{i\in \mathrm{I}}$ of $\mathbf{GL}_2(\mathbf{R})$, each occuring with finite multiplicities.

We also have:

$$V_\pi^{K_1}(\rho_\infty) = \widehat{\bigoplus_{i\in \mathrm{I}}} V_i(\rho_\infty)$$

so it suffices to prove that only finitely many i are such that $V_i(\rho_\infty) \neq \{0\}$.

One the other hand, $\mathfrak{z}$ acts by characters on the smooth vectors of V_i^∞ for each $i \in \mathrm{I}$: in particular, there exists a complex number λ such that $V := \ker(\Delta - \lambda \mathrm{Id}) \neq \{0\}$ (here we take the kernel in V_π^∞), where Δ denotes the Casimir element of $\mathbf{GL}_2(\mathbf{R})$. This subspace V is obviously stable under the action of $\mathbf{GL}_2(\mathbb{A}_f)$ and of $\mathbf{GL}_2(\mathbf{R})$: therefore it must be dense in V_π. This implies that Δ acts by λ on *each* V_i^∞ (by using the self-adjointness of Δ). By the classification of irreducible unitary representations of $\mathbf{GL}_2(\mathbf{R})$, there are only finitely many equivalence classes of representations of $\mathbf{GL}_2(\mathbf{R})$ containing ρ_∞, with central character ω_∞, such that Δ acts by λ on the smooth vectors: this ends the proof.

Remark: This argument works in generality for reductive groups (use [Bo2] théorème 5.29), but instead of using the Casimir element, one can argue as follows – this affects the last paragraph of the previous proof: $\mathfrak{z}$ acts by characters on the smooth vectors of V_i^∞ for each $i \in \mathrm{I}$: let χ be one of them. Let V be the space of smooth vectors son which $\mathfrak{z}$ acts through χ: V is dense in V_π for the same reasons. This implies that for *any* $v \in V_\pi$, $d\pi(z) = \chi(z)v$ as a distribution, for any $z \in \mathfrak{z}$ (by taking a sequence in V tending to v), and so that on *any* smooth vector of V_π, $\mathfrak{z}$ acts through χ. This implies that the infinitesimal character of each V_i is χ, and again, there are only finitely many irreducible unitary representations containing ρ_∞, with central character ω_∞ and infinitesimal character χ.

7. Proof of theorem 5.1

Let f be a cuspidal modular form of weight k, nebentypus χ and level N. Assume that there exists a finite set of primes S_f such that f is a Hecke eigenform for the T_p, $p \notin S_f$: we may and will assume that S_f contains the divisors of q. We wish to prove that that the unitary representation $(\mathfrak{H}_f, \pi_f)$ we attached to f in section 5 is irreducible.

To simplify the proof, we'll make use of the strong multiplicity one theorem, which asserts that given two irreducible unitary cuspidal representations $\pi \cong \widehat{\bigotimes_{v\leq\infty}} \pi_v, \pi' \cong \widehat{\bigotimes_{v\leq\infty}} \pi'_v$ of $\mathbf{GL}_2(\mathbb{A}_\mathbf{Q})$ are equivalent if and only if there exists a finite set of primes S (containing or not ∞) such that $\pi_v \cong \pi'_v$ for each $v \notin S$. This theorem holds actually for irreducible automorphic representations, and can be proven using Whittaker models (cf [Bu] section 3.5, [Go] §3.5 and [G] §6) or the Rankin-Selberg L-function (as in [CKM] theorem 9.3): in any case, the proof makes use of the algebraic theory of automorphic forms, in the sense that Whittaker models are smooth models, not unitary representations.

As $L^2_0(Z(\mathbb{A})\mathbf{GL}_2(\mathbf{Q})\backslash\mathbf{GL}_2(\mathbb{A}_\mathbf{Q}), \omega_\chi)$ is $\mathbf{GL}_2(\mathbb{A}_\mathbf{Q})$-invariant, $\mathfrak{H}_f$ is a subspace of it. Also, as $L^2_0(Z(\mathbb{A})\mathbf{GL}_2(\mathbf{Q})\backslash\mathbf{GL}_2(\mathbb{A}_\mathbf{Q}), \omega_\chi)$ splits as a Hilbert sums of irreducibles, then so does $(\mathfrak{H}_f, \pi_f)$ by Lemma 3.1. We may therefore write:

$$\mathfrak{H}_f = \widehat{\bigoplus_{i\in \mathrm{I}}} \mathfrak{H}_i$$

where each $\mathfrak{H}_i$ is a closed subspace of $\mathfrak{H}_f$ stable and irreducible under $\mathbf{GL}_2(\mathbb{A}_\mathbf{Q})$. We have to prove that $\mathrm{card}(\mathrm{I}) = 1$.

To do so, let's denote π_i the representation of $\mathbf{GL}_2(\mathbb{A}_\mathbf{Q})$ on $\mathfrak{H}_i$. By the tensor product theorem, for each i we can write (with alleged notations)

$$\pi_i \cong \widehat{\bigotimes_{v\leq\infty}} \pi_{i,v}.$$

To prove the theorem, due to the multiplicity one theorem, it is sufficent to prove that $\pi_{i,p} \cong \pi_{j,p}$ for any $p \notin S_f$ and each $i, j \in \mathrm{I}$.

As the adelization of f, φ_f, is in $\mathfrak{H}_f$, we can write, in a unique way – with convergence in L^2:

$$\varphi_f = \sum_{i\in \mathrm{I}} \varphi_i$$

with $\varphi_i \in \mathfrak{H}_i - \{0\}$. If K' denotes the product of the $\mathbf{GL}_2(\mathbf{Z}_p)$'s for $p \notin S_f$, then φ_f is K' right invariant, and we can assume that so are the φ_i's (if not, one writes $\varphi_f(g) = \frac{1}{\mathrm{vol}(K')} \int_{K'} \psi(k)\varphi_f(gk)dk = \sum_i \int_{K'} \psi(k)\varphi_i(gk)dk$ with ψ = the characteristic function of K', and the job is done, or one simply projects on the K'-invariant vectors).

Let $i \in \mathrm{I}$. By the tensor product theorem, one can write $\mathfrak{H}_i = \widehat{\bigotimes_{v \leq \infty}} \mathfrak{H}_{i,v}$, and so there exists a set J such that:

$$\varphi_i = \sum_{j \in J} x_j$$

where x_j is of the form $x_j = j_{S_j}(x_{S_j})$ with $x_{S_j} \in \mathfrak{H}_{i,S_j}$ – see section 6 for the notations. As we did above, one can assume that each of the x_j's are K'-invariant, i.e. that $S_j \subset S_f$. All this proves that the π_i's are unramified outside S_f.

Let $p \notin S_f$ be a prime. Obviously, all the π_i's have the same central character (namely ω): so do all the $\pi_{i,p}$ (for various i's). To prove the theorem, it is sufficient to prove that the $\pi_{i,p}$ share the same eigenvalue under the Hecke operator T_p we introduced in section 4.2. We defined in section 5.3 the adelization of the classical Hecke operator $\mathbb{T}_p$. By hypothesis, we have:

$$\mathbb{T}_p \varphi_f = \lambda_f(p) \varphi_f$$

and so, by continuity for each $i \in \mathrm{I}$:

$$\mathbb{T}_p \varphi_i = \lambda_f(p) \varphi_i. \tag{3}$$

As each x_j is a "pure tensor", say $x_j = \otimes_v x_{j,v}$ (for $p \notin S_f$, $x_{j,p} = \xi^0_{j,p}$, the K_p-fixed vector), one has :

$$\mathbb{T}_p x_j = \bigotimes_{v \neq p} x_{j,v} \otimes (T_p(x_{j,p})).$$

The vector $x_{j,p}$ is K_p-invariant, so by denoting $c_{i,p}$ the Hecke eigenvalue of $\pi_{i,p}$:

$$T_p(x_{j,p}) = c_{i,p} x_{j,p}.$$

By comparing with (3), one gets $\lambda_f(p) = c_{i,p}$ for any $i \in I$, so we get exactly what we wanted. QED

8. Appendix

8.1. **Appendix 1.** Let G be a locally compact unimodular group, H a closed unimodular subgroup of G, Z the center of G (or more generally a closed subgroup of the center) and $\omega : Z \to \mathbb{S}^1$ a character. We want to first define in this appendix what is meant in the literature by $L^2(ZH\backslash G, \omega)$, as it was mentioned in section 5.2. So let $L^2(ZH\backslash G, \omega)$ ($L^2(\omega)$ is a useful abbreviation if no confusion arises) be the space of classes of functions (the equivalence is equality almost everywhere on $H\backslash G$) $f : H\backslash G \to \mathbf{C}$ such that:

(1) f is Borel-measureable on $H\backslash G$,
(2) $|f|$ is Borel-measureable on $ZH\backslash G$,
(3) for any $z \in Z$, $f(zx) = \omega(z) f(x)$ for almost all $x \in H\backslash G$,
(4) $\int_{ZH\backslash G} |f|^2 < \infty$

We claim that:

(1) $L^2(ZH\backslash G,\omega)$ is a Hilbert space for the inner product $\langle\cdot,\cdot\rangle$ defined by:

$$\langle f,g\rangle=\int_{ZH\backslash G}f(x)\overline{g(x)}dx$$

(2) The space of bounded continuous functions $\mathcal{C}_b(ZH\backslash G,\omega)$ satisfying:

$$f(zx)=\omega(z)f(x)$$

for all $x\in H\backslash G$, is dense in $L^2(ZH\backslash G,\omega)$

Let (f_n) be a Cauchy sequence in $L^2(ZH\backslash G,\omega)$. One can find a subsequence φ_n such that:

$$\|\varphi_{n+1}-\varphi_n\|\leq 2^{-n}$$

This implies that the series $\sum_{n\in\mathbf{N}}|\varphi_{n+1}-\varphi_n|$ converges almost everywhere on $ZH\backslash G$, and thus that $\sum_{n\in\mathbf{N}}(\varphi_{n+1}-\varphi_n)$ is absolutely convergent almost everywhere on $H\backslash G$: this proves the first claim.

To prove the second claim, note first that the subspace $L^2_c(ZH\backslash G,\omega)$ of functions with compact support in $ZH\backslash G$ is dense in $L^2(ZH\backslash G,\omega)$: for instance, if K is a (large) compact subset of $ZH\backslash G$, $f\times\mathrm{Char}_{K'}$ will do the job (K' is the inverse image of K under the projection $H\backslash G\to ZH\backslash G$). Then, let f be in $L^2_c(ZH\backslash G,\omega)$, and let φ_n a continuous δ-sequence in G. Consider the function:

$$f_n(x)=\int_G f(xg)\varphi_n(g)dg$$

It is well-defined: denoting by K_n a compact of G containing the support of φ_n, one has:

$$\begin{aligned}\int_{ZH\backslash G}|f_n(x)|^2dx&\leq\int_{ZH\backslash G}\left(\int_G|\varphi_n(g)|^2dg\int_G|f(xg)|^2\mathrm{Char}_{K_n}(g)dg\right)dx\\&\leq\mathrm{vol}(K_n)\times\int_G|\varphi_n(g)|^2dg\times\int_{ZH\backslash G}|f(x)|^2dx\end{aligned}$$

The function f_n is continuous, as this is easily seen after a legal change of variable and a use of Lebesgue dominated convergence theorem, and its support is compact (because it is the convolution of two such functions). This ends our contention.

8.2. **On cuspidal functions.** We refer to [L] and [Bo1] (especially chap 8) for a rigorous definition of this space, in the classical setting. In the litterature, given a unitary Grossencharakter $\omega:\mathbb{A}_{\mathbf{Q}}^{\times}/\mathbf{Q}^{\times}\to\mathbf{C}^{\times}$, the space

of square-integrable cusp forms is often "defined" by:

$$L_0^2(Z(\mathbb{A}_{\mathbf{Q}})\mathbf{GL}_2(\mathbf{Q})\backslash\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}}),\omega) =$$
$$\Big\{ f : \mathbf{GL}_2(\mathbf{Q})\backslash\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}}) \to \mathbf{C} \ : \ \text{for all } z \in Z(\mathbb{A}_Q), f(zg) = \omega(z)f(g)$$
$$\text{for almost all } g, \int_{Z(\mathbb{A}_{\mathbf{Q}})\mathbf{GL}_2(\mathbf{Q})\backslash\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})} |f|^2 < \infty \text{ and}$$
$$\int_{\mathbf{Q}\backslash\mathbb{A}_{\mathbf{Q}}} f\left(\left[\begin{smallmatrix}1 & x\\ 0 & 1\end{smallmatrix}\right] g\right) dx = 0 \text{ for almost all } g \Big\}$$

The problem is that the last condition is not a closed one, a priori. One option is to define this space as the closure in $L^2(\omega)$ –which we defined in the previous paragraph – of the space of bounded continuous satisfying the above (well defined) conditions: this is what Lang does.

Another possibility is to consider, for a compactly supported function φ on $U(\mathbb{A}_{\mathbf{Q}})\backslash\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$ (here U denotes the usual unipotent subgroup of $\mathbf{GL}_2$), the linear form:

$$f \in L^2(\omega) \to \Lambda_\varphi(f) = \int_{U(\mathbf{Q})\backslash\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})} f(g)\varphi(g)dg$$

This mapping is well-defined, and continuous: indeed, the support of φ, viewed as a function on $\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$, is contained in a set of the shape $U(\mathbf{Q})\Omega$, where Ω is a compact subset of $\mathbf{GL}_2(\mathbb{A})$ – this is because $\mathbf{Q}$ is cocompact inside $\mathbb{A}_{\mathbf{Q}}$. So one has:

$$|\Lambda_\varphi(f)| \leq \|\varphi\|_\infty \int_\Omega |f(x)|dx$$

To finish, one covers Ω with finitely many (say m) relatively compact open sets U_i of $\mathbf{GL}_2(\mathbb{A})$ such that the projection $\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}}) \to \mathbf{GL}_2(\mathbf{Q})\backslash\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$ induces on each U_i a homeomorphism onto its image – which is possible by the discreteness of $\mathbf{GL}_2(\mathbf{Q})$ in $\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})$:

$$\int_\Omega |f(x)|dx \leq m \int_{Z(\mathbb{A}_{\mathbf{Q}})\mathbf{GL}_2(\mathbf{Q})\backslash\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})} |f(x)|dx \times \sup_i(\text{vol})(U_i)$$

this proves the continuity, because $\text{vol}(Z(\mathbb{A}_{\mathbf{Q}})\mathbf{GL}_2(\mathbf{Q})\backslash\mathbf{GL}_2(\mathbb{A}_{\mathbf{Q}})) < \infty$.

Finally, to see the link with the cuspidal condition, one notes that:

$$\Lambda_\varphi(f) = \int_{U(\mathbb{A}_{\mathbf{Q}})\backslash\mathbf{GL}_2(\mathbb{A})} W_f(g)\varphi(g)dg$$

where

$$W_f(g) := \int_{\mathbf{Q}\backslash\mathbb{A}_{\mathbf{Q}}} f\left(\left[\begin{smallmatrix}1 & x\\ 0 & 1\end{smallmatrix}\right] g\right) dx,$$

and the space of cusp forms can be identified with the intersection of the kernels of all the Λ_φ, when φ varies among such functions (it is a posteriori easy to see this, by using convolution with a δ-sequence).

References

[BCSGKK] Bump, D.; Cogdell, J. W.; de Shalit, E.; Gaitsgory, D.; Kowalsk i, E.; Kudla, S. S.: An introduction to the Langlands program. Lectures presented at the Hebrew University of Jerusalem, Jerusalem, March 12–16, 2001. Edited by Joseph Bernstein and Stephen Gelbart. Birkhuser Boston, Inc., Boston, MA, 2003. x+281 pp.

[BH] Bushnell, Colin J.; Henniart, Guy: The local Langlands conjecture for $\mathbf{GL}(2)$. Grundlehren der Mathematischen Wissenschaften, 335. Springer-Verlag, Berlin, 2006.

[Bo1] Borel, Armand: Automorphic forms on $\mathbf{SL}_2(\mathbf{R})$. Cambridge Tracts in Mathematics, 130. Cambridge University Press, Cambridge, 1997. x+192 pp.

[Bo2] Borel, Armand: Représentations de groupes localement compacts. Lecture Notes in Mathematics, Vol. 276. Springer-Verlag, Berlin-New York, 1972. iv+98 pp.

[Bour-Int] Bourbaki, Nicolas: Éléments de mathématique. Fascicule XXIX. Livre VI: Intégration. Chapitre 7: Mesure de Haar. Chapitre 8: Convolution et représentations. Actualités Scientifiques et Industrielles, No. 1306 Hermann, Paris 1963 222 pp

[Bour-EVT] Bourbaki, Nicolas: Espaces vectoriels topologiques. Chapitres 1 5. Éléments de mathématique. Masson, Paris, 1981. vii+368 pp.

[BR] Barut, Asim O.; Raczka, Ryszard: Theory of group representations and applications. Second edition. World Scientific Publishing Co., Singapore, 1986. xx+717 pp.

[Bu] Bump, Daniel: Automorphic Forms and Representations, Cambridge Studies in Advanced Mathematics, 55, (1997).

[C] Casselman, William: On some results of Atkin and Lehner. Math. Ann. 20 1 (1973), 301–314.

[CF] Algebraic number theory. Edited by J. W. S. Cassels and A. Frhlich Academic Press, London; Thompson Book Co., Inc., Washington, D.C. 1967 xviii+366 pp.

[CKM] Cogdell, James W.; Kim, Henry H.; Murty, M. Ram: Lectures on automorphic L-functions. Fields Institute Monographs, 20. American Mathematical Society, Providence, RI, 2004. xii+283 pp.

[Cor] Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 1, Amer. Math. Soc., Providence, R.I., 1979

[CSS] Modular forms and Fermat's last theorem. Edited by Gary Cornell, Joseph H. Silverman and Glenn Stevens. Springer-Verlag, New York, 1997. xx+582 pp.

[Dix] Dixmier, Jacques: C^*-algebras. Translated from the French by Francis Jellett. North-Holland Mathematical Library, Vol. 15. North-Holland Publishing Co., Amsterdam-New York-Oxford, 1977. xiii+492 pp.

[G] Gelbart, Stephen S.: Automorphic forms on adele groups, Annals of Math. studies, 83, Princeton University Press (1975).

[Ga1] Garrett, Paul: Proving Admissibility of Irreducible Unitaries, online at `http://www.math.umn.edu/~garrett/m/v/proving_admissibility.pdf`

[Ga2] Garrett, Paul: Factorization of unitary representations of adele groups, online at `http://www.math.umn.edu/~garrett/m/v/factoring_repns.pdf`

[Go] Godement, Roger: Notes on Jacquet-Langlands' theory, IAS lecture notes, 1970.

[Gui] Guichardet, A. Produits tensoriels infinis et reprsentations des relations d'anticommutation. (French) Ann. Sci. cole Norm. Sup. (3) 83 1966 1–52.

[GGP] Gel'fand, I. M.; Graev, M. I.; Pyatetskii-Shapiro, I. I.: Representatio n theory and automorphic functions. Translated from the Russian by K. A. Hirsch W. B. Saunders Co., Philadelphia, Pa.-London-Toronto, Ont. 1969 xvi+426 pp.

[I] Iwaniec, Henryk: Topics in classical automorphic forms. Graduate Studies in Mathematics, 17. American Mathematical Society, Providence, RI, 1997. xii+259 pp.

[IK] H. Iwaniec, E. Kowalski: Analytic Number Theory, American Mathematical Society Colloquium Publications, 53. American Mathematical Society, Providence, RI, 2004. xii+615 pp.

[ILS] Iwaniec, Henryk; Luo, Wenzhi; Sarnak, Peter: Low lying zeros of families of L-functions. Inst. Hautes tudes Sci. Publ. Math. No. 91 (2000), 55–131 (2001).

[KL] Knightly, Andrew; Li, Charles: Traces of Hecke operators. Mathematical Surveys and Monographs, 133. American Mathematical Society, Providence, RI, 2006. x+378 pp.

[L] Lang, Serge: $\mathbf{SL}_2(\mathbf{R})$. Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, 1975. xvi+428 pp.

[M] Miyake, Toshitsune: Modular forms. Translated from the Japanese by Yoshitaka Maeda. Springer-Verlag, Berlin, 1989. x+335 pp.

[M2] Miyake, Toshitsune On automorphic forms on GL_2 and Hecke operators. Ann. of Math. (2) 94 (1971), 174–189.

[Ro] Robert, Alain: Introduction to the representation theory of compact and locally compact groups. London Mathematical Society Lecture Note Series, 80. Cambridge University Press, Cambridge-New York, 1983. ix+205 pp.

[Su] Sugiura, Mitsuo Unitary representations and harmonic analysis. An introduction. 2nd ed. North-Holland Mathematical Library, 44. North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, 1990. xvi+452 pp.

[T] Taylor, Richard: Galois representations. Ann. Fac. Sci. Toulouse Math. (6) 13 (2004), no. 1, 73–119.

[War] Warner, Garth: Harmonic analysis on semi-simple Lie groups. I and II. Die Grundlehren der mathematischen Wissenschaften, Band 188/189. Springer-Verlag, New York-Heidelberg, 1972.

[Wal] Wallach, Nolan R.: Real reductive groups. I and II. Pure and Applied Mathematics, 132 and 132-II. Academic Press, Inc., Boston, MA, 1988 and 1992.

[Wal1] Wallach, Nolan R. C^∞ vectors. Representations of Lie groups and quantum groups (Trento, 1993), 205–270, Pitman Res. Notes Math. Ser., 311, Longman Sci. Tech., Harlow, 1994.

[We] Weil, André: Dirichlet Series and automorphic forms, Lecture Notes in Mathematics, Vol. 189 Springer-Verlag, Berlin-New York, 1971.

Stanford University, Department of Mathematics, building 380, Stanford, California 94305, USA. E-mail: `trotabas@math.stanford.edu`

Lecture 11: Hecke characters and Galois characters

Andrew Snowden
January 28, 2010

1. Introduction

Let K be a number field, let $\mathbf{A}_K^\times$ be its idele group and let G_K be its absolute Galois group. Class field theory states that there is a natural map (up to a choice of normalization)

$$\mathbf{A}_K^\times / K^\times \to G_K^{\mathrm{ab}}$$

which identifies G_K^{ab} as the profinite completion of $\mathbf{A}_K^\times/K^\times$. Equivalently, class field theory can be stated as an isomorphism

$$\{\text{finite order characters of } \mathbf{A}_K^\times/K^\times\} = \{\text{finite order characters of } G_K\}.$$

Thus we have a description of the finite order characters of G_K.

A *p-adic character* of G_K is a continuous homomorphism $G_K \to \overline{\mathbf{Q}}_p^\times$; since G_K is compact any such character takes values in $\mathscr{O}_F^\times$ for some finite extension $F/\mathbf{Q}_p$. There are many interesting p-adic characters which are not of finite order: for instance, the cyclotomic character χ_p. Since $\mathscr{O}_F^\times$ is profinite, p-adic characters of G_K are limits of finite order characters, and so we can use class field theory to understand them. Define a *p-adic Hecke character* (of K) to be a continuous homomorphism $\mathbf{A}_K^\times/K^\times \to \overline{\mathbf{Q}}_p^\times$; again, the image is always contained in $\mathscr{O}_F^\times$ for some F finite over $\mathbf{Q}_p$. We then have an identification

$$\{p\text{-adic Hecke characters of } K\} = \{p\text{-adic characters of } G_K\}.$$

induced by class field theory.

We thus have an understanding of p-adic characters of the Galois group. However, this is not the end of the story: there are *compatible systems* of characters. Such a system consists of a p-adic character ψ_p of G_K for each prime p such that for each place v of K the quantity $\psi_p(\mathrm{Frob}_v)$ is independent of p in a suitable sense. We would like to understand the collection of compatible systems. The Langlands program suggests that compatible systems of characters should correspond to automorphic representations of $\mathrm{GL}_1(\mathbf{A}_K)$, so we now examine these objects.

What is an automorphic representation of $\mathrm{GL}_1(K)$? To begin with, it should be an irreducible subrepresentation of $\mathrm{GL}_1(\mathbf{A}_K)$ acting on the space of automorphic forms on $\mathrm{GL}_1(\mathbf{A}_K)$ by right translation. (Recall that an automorphic form on $\mathrm{GL}_1(K)$ is a function $K^\times\backslash \mathrm{GL}_1(K) \to \mathbf{C}$ satisfying certain natural conditions.) Since $\mathrm{GL}_1(\mathbf{A}_K)$ is commutative, such a representation must be one dimensional. It is thus spanned by some non-zero automorphic form f. Since $\mathbf{C}f$ is stable by right translation, we find $f(xg) = \lambda_g f(x)$ for all $x, g \in \mathrm{GL}_1(\mathbf{A}_K)$. Taking $x = 1$, we find $\lambda_g f(1) = f(g)$ and so $f(1)f(xg) = f(g)f(x)$ holds for all x and g. Since f is non-zero we find that $f(1)$ is non-zero; scale f so that $f(1) = 1$. We then find that f is a homomorphism, and since it is invariant under $K^\times$ it also satisfies $f(K^\times) = 1$. The properties of automorphic forms that we did not list amount to f being continuous. A continuous homomorphism $\mathbf{A}_K^\times/K^\times \to \mathbf{C}^\times$ is called a *Hecke character*. We have thus shown that every automorphic representation of $\mathrm{GL}_1(K)$ is spanned by a Hecke character. It is clear that the character is unique. It is also not difficult to show that every Hecke character spans an automorphic representation. We thus have an identification

$$\{\text{automorphic representations of } \mathrm{GL}_1(K)\} = \{\text{Hecke characters of } K\}.$$

Consider now the diagram

$$\begin{array}{ccc} \{\text{Hecke characters of } K\} & \dashrightarrow & \{\text{compatible systems of characters of } G_K\} \\ \Big\downarrow & & \Big\downarrow \\ \{p\text{-adic Hecke characters of } K\} & = & \{p\text{-adic characters of } G_K\} \end{array}$$

We have already explained the bottom map. The right map takes a compatible system of characters $\{\psi_p\}$ to its pth member ψ_p. The top arrow means "we expect a relationship." Given a top arrow, the left arrow is obtained by going around the diagram.

As we have introduced it, the left arrow might seem the most mysterious: it is given by combining three operations, one of which is itself somewhat unclear. However, it is actually quite accessible. We will explain this in the coming sections. Once one has this left arrow, it is not difficult to understand the top arrow. The story goes like this. There are certain special Hecke characters, the *algebraic* ones. Given an algebraic Hecke character f one can build a p-adic Hecke character f_p for any prime p. Each f_p is associated to some p-adic character ψ_p of G_K and these ψ_p form a compatible system. In fact, this is a bijection, that is, every compatible system arises from a unique algebraic Hecke character.

2. The case $K = \mathbf{Q}$

We begin by considering the case $K = \mathbf{Q}$. The general case does not differ much from this case except that it is more notationally complicated. We have

$$\mathbf{A}_{\mathbf{Q}}^{\times}/\mathbf{Q}^{\times} = \prod_p \mathbf{Z}_p^{\times} \times \mathbf{R}_+.$$

Here $\mathbf{R}_+$ denotes the group of positive real numbers under multiplication. Each $\mathbf{Z}_p^{\times}$ has its usual topology and the product has the product topology; it is profinite.

Let f be a Hecke character. The restriction η of f to $\prod \mathbf{Z}_p^{\times}$ is a finite order character, as is any continuous homomorphism from a profinite group to $\mathbf{C}^{\times}$. The restriction of f to $\mathbf{R}_+$ is of the form $x \mapsto x^a$ for some real number a. We call f *algebraic* if this number a is an integer. Let α_∞ be the Hecke character which is trivial on $\prod \mathbf{Z}_p^{\times}$ and on $\mathbf{R}_+$ is given by the standard inclusion $\mathbf{R}_+ \to \mathbf{C}^{\times}$. Then an arbitrary Hecke character f is algebraic if and only if it is of the form $\eta\alpha_\infty^n$ for some finite order character η and integer n. The character η and the integer n are then uniquely determined.

Now let f be a p-adic Hecke character. The restriction of f to $\mathbf{R}_+$ is then trivial. The restriction of f to $\prod_{\ell \neq p} \mathbf{Z}_\ell^{\times}$ is of finite order. The restriction of f to $\mathbf{Z}_p^{\times}$ is a continuous homomorphism $\mathbf{Z}_p^{\times} \to \overline{\mathbf{Q}}_p^{\times}$. It is not difficult to classify all such maps, but we will not do this. We call f *algebraic* if this restriction is of the form $x \mapsto x^n$ on a compact open subset of $\mathbf{Z}_p^{\times}$. Let α_p be the p-adic Hecke character which is trivial on $\mathbf{R}_+$ and $\prod \mathbf{Z}_\ell^{\times}$ and on $\mathbf{Z}_p^{\times}$ is given by the standard inclusion $\mathbf{Z}_p^{\times} \to \overline{\mathbf{Q}}_p^{\times}$. Then an arbitrary p-adic Hecke character f is algebraic if and only if it is of the form $\eta\alpha_p^n$ for some finite order character η and integer n. Again, η and n are uniquely determined.

Let f be an algebraic Hecke character. We can then write $f = \eta\alpha_\infty^n$. Define a p-adic Hecke character f_p by $f_p = \eta\alpha_p^n$. (Here we are implicitly identifying the roots of unity in $\mathbf{C}$ and $\overline{\mathbf{Q}}_p$ so that we may regard η as being valued in either field.) Under class field theory, the p-adic Hecke character α_p corresponds to the cyclotomic character χ_p. Thus f_p corresponds to $\psi_p = \eta'\chi_p^n$, where η' is the finite order character of G_K corresponding to η. Since the χ_p form a compatible system, we thus see that the ψ_p do as well. Therefore, starting from a Hecke character f we can produce a system $\{f_p\}$ of p-adic Hecke characters and from this obtain a compatible system $\{\psi_p\}$ of one dimensional Galois representations.

3. The general case

Let K be an arbitrary number field. We will find it convenient to treat p-adic Hecke characters and normal Hecke characters (which we now call ∞-adic Hecke characters) simultaneously. Thus let p be a prime or ∞. Let C_p be $\overline{\mathbf{Q}}_p$ or $\mathbf{C}$ correspondingly (one could use $\mathbf{C}_p$ in place of $\overline{\mathbf{Q}}_p$). A p-adic Hecke character is then just a continuous homomorphism

$$\mathbf{A}_K^{\times}/K^{\times} \to C_p^{\times}.$$

We fix an emedding $K \to C_p$ for each p. We explain how this large number of choices can be cut down at the end of the section.

Let f be a p-adic Hecke character. We regard f as a character of

$$\mathbf{A}_K^{\times} = (K \otimes \mathbf{Q}_p)^{\times} \times \prod_{\ell \neq p} (K \otimes \mathbf{Q}_\ell)^{\times}$$

which is invariant under $K^{\times}$. (If $p = \infty$ then $\mathbf{Q}_p$ means $\mathbf{R}$.) Of course, f restricts to a finite order character on the second factor since ℓ-adic and p-adic topologies do not interact. On the first factor, however, f can be much more complicated. We say that f is *algebraic* if its restriction to the first factor is given by a rational function on an open subgroup, in the following sense. Regard $K \otimes \mathbf{Q}_p$ as an n-dimensional $\mathbf{Q}_p$ vector space, where $n = [K : \mathbf{Q}_p]$, and let $x_i : K \otimes \mathbf{Q}_p \to \mathbf{Q}_p$ be the n coordinates. Then we want f to be

a rational function function of the x_i, with coefficients in $\overline{\mathbf{Q}}_p$, after it is restricted to some open subgroup of $(K \otimes \mathbf{Q}_p)^\times$. (Note that we do not say compact open here. If $K \otimes \mathbf{Q}_p = \mathbf{R}$ then we allow f to be the absolute value character. This is an algebraic function when restricted to the open subgroup $\mathbf{R}_+$.)

We now give a nicer reformulation of this algebraic condition. Define a *weight* to be a character of the torus $\mathrm{Res}^K_{\mathbf{Q}}(\mathbf{G}_m)$ over $\overline{\mathbf{Q}}$, that is, a weight is a homomorphism of algebraic groups

$$w : (\mathrm{Res}^K_{\mathbf{Q}}(\mathbf{G}_m))_{\overline{\mathbf{Q}}} \to (\mathbf{G}_m)_{\overline{\mathbf{Q}}}.$$

Here Res denotes restriction of scalars. A weight gives a homomorphism $K^\times \to \overline{\mathbf{Q}}^\times$ which we also call w. For f to be algebraic, it is equivalent to ask that there exists a weight w and an open subgroup U of $(K \otimes \mathbf{Q}_p)^\times$ such that $f(x) = w(x)$ for $x \in K^\times \cap U$. The weight w is then unique and called the weight of f.

Before continuing we give an example. Let $K = \mathbf{Q}(\sqrt{d})$ be an imaginary quadratic field. Let u be a root of d in K and let u' be a root of d in $\mathbf{C}$. We fix an embedding $K \to \mathbf{C}$ by $u \mapsto u'$. Let f be a (∞-adic) Hecke character. At infinity, f gives a map $(K \otimes \mathbf{R})^\times \to \mathbf{C}^\times$. Now, $(K \otimes \mathbf{R})^\times$ is isomorphic to $\mathbf{C}^\times$. Every homomorphism $\mathbf{C}^\times \to \mathbf{C}^\times$ is of the form

$$re^{i\theta} \mapsto r^a e^{in\theta} = e^{a \log r + in\theta}$$

where a is a complex number and n is an integer. We wish to rephrase this classification in terms of K. Every element of K can be written as $x + yu$ with x and y in $\mathbf{Q}$. In these coordinates, r is given by the positive square root of $x^2 + dy^2$ while $e^{i\theta}$ is $(x + du')/\sqrt{x^2 + dy^2}$. We thus find that every homomorphism $(K \otimes \mathbf{R})^\times \to \mathbf{C}^\times$ is of the form

$$x + yu \mapsto (x^2 + dy^2)^{a/2} \left(\frac{x + du'}{\sqrt{x^2 + dy^2}} \right)^n = (x^2 + dy^2)^{(a-n)/2} (x + du')^n$$

where a is a complex number and n is an integer. Of course, a and n are uniquely determined. We find that this is a rational function of x and y if and only if $a - n$ is an even integer, say $2m$. In this case, the above formula can be written as

$$x + yu \mapsto (x - du')^m (x + du')^{n+m}.$$

Thus if we identify $K \otimes \mathbf{R}$ with $\mathbf{C}$ via $u \mapsto u'$ then any algebraic character $(K \otimes \mathbf{R})^\times \to \mathbf{C}^\times$ is of the form $z \mapsto z^n \overline{z}^m$. Of course, n and m are uniquely determined, but if we use the embedding $u \mapsto -u'$ then n and m are switched.

We now examine the same example from the point of view of weights. Let T be the torus $\mathrm{Res}^K_{\mathbf{Q}}(\mathbf{G}_m)$. It can be thought of as the group of all matrices inside of $\mathrm{GL}_2(\mathbf{Q})$ of the form

$$\begin{pmatrix} x & y \\ dy & x \end{pmatrix}$$

The group of characters $T_{\overline{\mathbf{Q}}} \to (\mathbf{G}_m)_{\overline{\mathbf{Q}}}$ is a free abelian group of rank two, generated by the two characters

$$\begin{pmatrix} x & y \\ dy & x \end{pmatrix} \mapsto x + uy, \qquad \begin{pmatrix} x & y \\ dy & x \end{pmatrix} \mapsto x - ux.$$

We thus see that the maps $K^\times \to \overline{\mathbf{Q}}^\times$ coming from weights are exactly the ones of the form

$$(x + uy) \mapsto (x + uy)^n (x - uy)^m.$$

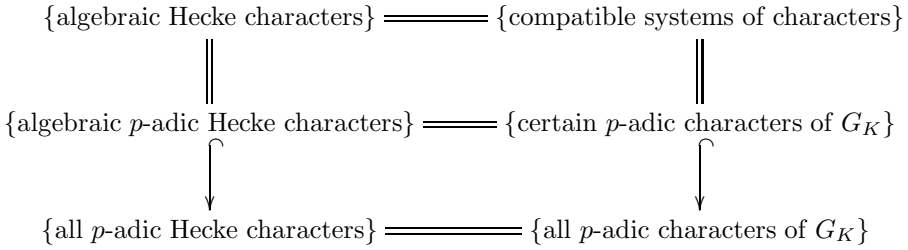
(Here we identify $T(\mathbf{Q})$ with $K^\times$ by letting $x + uy$ correspond to the matrix whose top left entry is x and bottom right entry is y.) Thus the homomorphisms $(K \otimes \mathbf{R})^\times \to \mathbf{C}^\times$ which come from weights are exactly the ones $z \mapsto z^n \overline{z}^m$. This shows that our two characterizations of algebraic homomorphisms agree in this case.

We now return to the general setting. Let f be an algebraic ∞-adic Hecke character with weight w. For any prime p let $\alpha_{w,p}$ be the homomorphism $\mathbf{A}_K^\times \to C_p^\times$ which is trivial on the prime to p components of $\mathbf{A}_K^\times$ and given by $w(x)$ for $x \in (K \otimes \mathbf{Q}_p)^\times$. Since f is algebraic of weight w, the character $\eta = f\alpha_{w,\infty}^{-1}$ is locally constant on $\mathbf{A}_K^\times$, that is, its kernel is open. It is not difficult to see that there is a number field M such that η takes values in $M^\times$. Choose an embedding of M into $\overline{\mathbf{Q}}_p$. We now define f_p to be $\eta\alpha_{w,p}$. We have thus associated a family of algebraic p-adic Hecke characters $\{f_p\}$ to our initial Hecke character f. Letting ψ_p be the character of G_K associated to f_p we also get a system $\{\psi_p\}$ of Galois characters. In fact, this is a compatible system.

Throughout we have been assuming an embedding of K into C_p for each p. This was actually not used in our first characterization of algebraic. However, it was used to move between f and f_p. We now explain a nicer set-up that requires fewer choices. The idea is to work with a field of coefficients. Let M be a number field. For a place v of M define a v-adic Hecke character to be a continuous homomorphism $\mathbf{A}_K^\times/K^\times \to M_v^\times$. Then, in this setting, one need only choose an embedding $K \to M$. That is, after having picked an embedding one can take an algebraic v-adic Hecke character and form from it a system of algebraic Hecke characters indexed by the places of M.

4. Conclusion

We have a diagram



We now explain this. Equal signs mean isomorphism. The top left vertical arrow is our construction to go between algebraic Hecke characters and algebraic p-adic Hecke characters. It is clearly an isomorphism, since one can run the construction in reverse. The bottom left vertical arrow is just the inclusion of the algebraic characters into all characters. The bottom horizontal arrow has already been discussed. The "certain p-adic characters of G_K" are just those that come from algebraic Hecke characters. The middle horizontal and bottom right vertical arrow are evident. The top right vertical arrow takes a compatible system to its pth member. It is injective since a compatible system is determined by any of its members. We now come to interesting part. We have shown how to attach to an algebraic Hecke character a system of p-adic Hecke characters, and therefore a system of p-adic characters of G_K. As we stated, this is a compatible system. This gives a map from the top left to the top right. It is easily seen to be injective. A more difficult result is that it is surjective — every compatible system is associated to an algebraic Hecke character. A diagram chase now gives the surjectivity of the top right vertical arrow.

A natural question one may now ask is: which are the "certain" p-adic characters of G_K that arise from algebraic Hecke characters? One answer is provided by the diagram: they are exactly those that fit into a compatible system of characters. There is a better answer, though, one that is intrinsic to the character. Namely, a p-adic character of G_K comes from an algberaic Hecke character if and only if it is Hodge-Tate. This is a condition from p-adic Hodge theory.

AUTOMORPHIC FORMS ON QUATERNION ALGEBRAS

ANDREW SNOWDEN

Let F be a totally real number field and let D be a quaternion algebra over F. Define an algebraic group G over F by $G(A) = (A \otimes_F D)^\times$ for an F-algebra A. It is easy to see that if F'/F splits D then $G_{F'}$ is isomorphic to $\mathrm{GL}(2)$. Thus G is a reductive algebraic group. We therefore have a theory of automorphic forms and representations for G. We will look at the basics of this theory in these notes.

To begin with, let us examine the spaces on which automorphic forms are functions. Let f be an automorphic form on $G(\mathbf{A}_F)$. By definition, f is a function $G(\mathbf{A}_F) \to \mathbf{C}$ subject to the following:

- f is left invariant under $G(F)$.
- f is invariant under a compact open subgroup U of $G(\mathbf{A}_{F,f})$.
- f is finite under translations by a maximal compact subgroup K of $G(F \otimes \mathbf{R})$.
- f is finite under the center of the universal enveloping algebra of $G(F \otimes \mathbf{R})$.
- f satisfies certain continuity and growth conditions.

For simplicity, let us consider the case where f is invariant under the center Z of $G(\mathbf{A}_F)$ and transforms under K by a one dimensional representation σ, i.e., $f(gk) = \sigma(k)f(g)$ holds. Then f defines a section of a line bundle determined by σ on the space

$$X(U) = G(F)\backslash G(\mathbf{A}_F)/ZKU.$$

Our first task is to describe this space.

The most important thing to initially consider about $X(U)$ is the contributions of the infinite places. At an infinite place v the division algebra D has two possibles behaviors: it can either split or not. In either case, $D_v^\times$ is a four dimensional real Lie group. If D_v is split then G_v looks like $\mathrm{GL}_2(\mathbf{R})$ and so its maximal compact is the two dimensional orthogonal group. We thus find that G_v/K_vZ_v is a copy of the upper half plane. In particular, it is a one dimensional complex manifold. If D_v is non-split then G_v is the multiplicative group in the Hamilton quaternions. This group is an extension of the rank 2 unitary group by $\mathbf{R}_+$. Thus $G_v = Z_vK_v$ and so the quotient G_v/Z_vK_v is a point. We therefore find

$$G(F \otimes \mathbf{R})/Z_\infty K = \mathfrak{h}^n$$

where n is the number of infinite places at which D is split. This computation is significant for two reasons. First, we see that the quotient is canonically a complex manifold, so we can make sense of holomorphic functions on it. And secondly, if $n = 0$, that is, if D is non-split at all infinite places, then this space is just a point.

Consider the case $n > 0$. Since $G(F \otimes \mathbf{R})$ is non-compact, strong approximation gives

$$G(\mathbf{A}_F) = G(F)G(F \otimes \mathbf{R})U$$

and so the usual computation show that

$$X(U) = \Gamma(U)\backslash\mathfrak{h}^n$$

where $\Gamma(U)$ is the arithmetic group obtained from intersecting $G(F)$ and U. (This assumes something about U: the norm map $U \to \prod \mathbf{G}_m(\mathscr{O}_{F,v})$ is surjective. For a general U, $X(U)$ will not be connected but will have finitely many connected components, each of the above form.) Thus $X(U)$ is a complex manifold obtained as the quotient of n copies of the upper half plane by the action of a totally discontinuous subgroup. When $n = 1$ (and $F \neq \mathbf{Q}$) these spaces are called *Shimura curves*. In contrast to the modular curves, they are compact — no cusps need to be added to obtain a compact space.

Now consider the case $n = 0$. Strong approprixmation no longer applies to $G(F \otimes \mathbf{R})$ since this group is compact. However, we do not really need to use strong approximation. Since KZ contains all of $G(F \otimes \mathbf{R})$

Date: January 28, 2010.

we can completely ignore the infinite places. We find

$$X(U) = G(F)\backslash G(\mathbf{A}_{F,f})/Z_f U.$$

Since U is a compact open subgroup of $G(\mathbf{A}_{F,f})$, this quotient is discrete. In fact, it is a finite set; this follows from it being discrete and having finite volume. Thus in this case, automorphic forms on D are simply functions on a finite set. There are therefore no continuity or analytic conditions placed on the forms — they are just functions on a finite set!

We did not consider the most general possible set-up. One does not need to assume Z-invariance: one can allow f to transform under Z by a character. Also, the representation σ of K, which plays the role of the weight, does not have to be one dimensional. (One must then take f to be vector valued.) The rank two unitary group has irreducible representations of all dimensions, so when D is non-split at infinite places one might want to consider such representations of K.

We now change directions and consider automorphic representations of $G(\mathbf{A}_F)$. Such a representation decomposes as a tensor product $\otimes\pi_v$ over the places of F, where π_v is an irreducible admissible representation of G_v. For almost all places, D_v is split, and so G_v is isomorphic to $\mathrm{GL}_2(F_v)$. The local invariants (such as conductor, L-series and ϵ-factors) at these places are defined as usual. We now consider a place v at which D_v is non-split. The group G_v is compact modulo its center. This implies that the representation π_v is finite dimensional. The maximal compact subgroup K_v of G_v has a natural filtration $K_v^{(n)}$ obtained by looking at the group of elements congruent to 1 modulo powers of the maximal ideal. There is a unique minimal n such that π_v contains $K_v^{(n)}$ in its kernel. The number $n+1$ is called the *conductor* of π_v (or perhaps the exponent of the conductor). The prime power $\mathfrak{p}_v^{n+1}$ is the local contribution of π_v to the level. Note that this exponent is never 0 — even when π_v contains the full maximal compact subgroup K_v in its kernel the conductor is 1. The reason for this will be evident later — suffice it to say for now that the Galois representations coming from modular forms on a quaternion algebra are always ramified where the quaternion algebra is, so these places should appear in the conductor.

We can additionally attach local L-functions and ϵ-factors to π_v. This goes quite similarly to the GL(1) case (Tate's thesis). For a Schwartz function ϕ on D_v one considers the integral

$$Z(s, \phi, \pi_v) = \int_{D_v} \phi(x) \operatorname{tr} \pi_v(x) |x|^{s+\cdots} dx.$$

The elipses in the exponent is a normalizing factor, which is not important for the present discussion. One then finds that there is a unique Euler factor $L(s, \pi_v)$ such that the quotient

$$\frac{Z(s, \phi, \pi_v)}{L(s, \pi_v)}$$

is an entire function of s and for some choice of ϕ is equal to 1. Furthermore, there is a functional equation

$$\frac{Z(1-s, \hat{\phi}, \pi_v^\vee)}{L(1-s, \pi_v^\vee)} = \epsilon(s, \pi_v, \psi_v) \frac{Z(s, \phi, \pi_v)}{L(s, \pi_v)}.$$

Here ψ_v is a non-trivial additive character of F_v, $\hat{\phi}$ denotes the Fourier transform of ϕ with respect to ψ_v and $\pi_v^\vee$ denotes the contragredient of π_v. The factor $\epsilon(s, \pi_v, \psi_v)$ is of the form $s \mapsto ab^s$. The base b of this exponential is equal (or almost equal) to the conductor of π_v.

Generalities on Central Simple Algebras

Michael Lipnowski

Introduction

The goal of this talk is to make readers (somewhat) comfortable with statements like "*the* quaternion algebra over $\mathbb{Q}$ ramified at 2,5,7,11." Statements like this will come up all the time, when we use Jacquet-Langlands.

The Basic Theorems

Definition 1. *A **central simple algebra** (CSA) over a field k is a finite dimensional k-algebra with center k and no non-trivial two-sided ideals.*

Some examples:

- Any division algebra over k is clearly a central simple algebra since any non-zero element is a unit. For example, we have quaternion algebras:

  $$H(a,b) = span_k\{1,i,j,ij\}$$

  with multiplication given by

  $$i^2 = a, j^2 = b, ij = -ji.$$

  For example, when $k = \mathbb{R}, a = b = -1$, we recover the familiar Hamilton quaternions $\mathbb{H}$.

- Let G be a finite group and $\rho : G \to GL_n(k)$ be an irreducible k-representation. Then $End_G(\rho)$ is a division algebra by Schur's Lemma. Hence, it is a CSA.

- $M_n(k)$ is a CSA. Indeed the left ideals are of the form

  $$\begin{pmatrix} * & 0 & * & 0 \\ * & 0 & * & 0 \\ * & 0 & * & 0 \\ * & 0 & * & 0 \end{pmatrix}$$

  and right ideals have a similar "transpose" shape.

A first step to understanding division algebras are the following basic theorems.

Double Centralizer Theorem 1. *Let A be a k-algebra and V a faithful, semi-simple A-module. Then*

$$C(C(A)) = End_k(V),$$

where the centralizers are taken in $End_k(V)$.

Classification of simple k-algebras. *Every simple k-algebra is isomorphic to $M_n(D)$ for some division k-algebra D.*

Proof. Choose a simple A-module S (for example, a minimal left ideal of A).
A acts faithfully on S since the kernel of $A \to End_k(S)$ is a two-sided ideal not containing 1.
Let D be the centralizer of A in the k-algebra $End_k(S)$. By the double centralizer theorem, $A = C(D)$, i.e. $A = End_D(S)$.
But S is a simple A-module. Thus for $d \in D$ multiplication by d is an A-linear endomorphism $d : S \to S$ and hence is either 0 or invertible, by Schur's Lemma. Since the inverse is also A-linear and $D = C(A)$, it follows that D is a division algebra.
It follows that D is a division k-algebra and so $S \cong D^n$ for some n. Hence,

$$A = End_D(D^n) = M_n(D^{opp}).$$

□

Noether-Skolem Theorem. *Let A be a simple k-algebra and B a semi-simple k-algebra. If*

$$f, g : A \to B$$

are k-algebra maps, then there is an invertible $b \in B$ such that

$$f(a) = bg(a)b^{-1}$$

for all $a \in A$.

The Brauer Group

We note the two following facts:

Proposition. *If A and B are CSAs over k, then $A \otimes_k B$ is a CSA over k.*

Proposition. *Let A be a CSA over k. Then $A \otimes_k A^{opp} \cong End_k(V)$.*

We define an equivalence relation on the set of CSAs over k by

$$A \sim B \text{if} A \otimes_k M_n(k) \cong B \otimes_k M_m(k) \text{for some positive integers} n, m.$$

This allows us to define the **Brauer group of** $k, Br(k)$, to be the set of equivalence classes of CSAs over k.
As of right now, this is only a Bruaer set. But we can endow it with a group operation $[A][B] = [A \otimes_k B]$.

- This operation is well-defined on equivalence classes since $M_n(k) \otimes_k M_m(k) \cong M_{mn}(k)$.
- It is clearly associative and commutative.
- $[k]$ acts as the identity.
- $[A^{opp}]$ is the inverse of $[A]$

Some examples of Baruer groups:

- If k is algebraically closed, then $Br(k) = 0$. Indeed, any CSA A is isomorphic to $M_n(D)$ for some division k-algebra D. But for any element $d \in D, k[d]$ is a finite field extension of k. But k is algebraically closed! Hence, $D = k$. But then $A \sim k$.
- $Br(R) = \{[\mathbb{R}], [\mathbb{H}]\}$.
- By a theorem of Wedderburn, all division algebras over finite fields are commutative. Hence, $Br(\text{finite field}) = 0$.

Extension of Base Field

Proposition. *Let A be a CSA over $k, K \supset k$ a field extension. Then $A \otimes_k K$ is a CSA over K.*

In this statement, K need not necessarily be finite over k.

Corollary. *If A is a CSA over k, then $[A : k]$ is a square.*

Proof.

$$[A : k] = [A \otimes_k \overline{k} : \overline{k}].$$

But $A \otimes_k \overline{k}$ is a CSA over $\overline{k}$ and so is isomorphic to $M_n(\overline{k})$ for some n. Thus, $[A : k] = n^2$. □

Note that if L/k is any field extension, then

$$Br(k) \to Br(K) : A \mapsto A \otimes_k L$$

defines a homomorphism. We let $Br(L/k)$ denote its kernel.

Brauer Groups and Cohomology

There is a natural isomorphism $H^2(L/k) = H^2(G_{L/k}, L^\times) \to Br(L/k)$. This is very handy, as it allows $Br(L/k)$ and $H^2(L/k)$ to play off each other. For example, it is not otherwise clear that $Br(k)$ is torsion or that $H^2(G_k, k^\times) = H^2(G_k^{un}, k^\times)$ for a local field k.
Here is a slightly more general version of the double centralizer theorem that we'll find useful.

Theorem. *Let B be a simple k-subalgebra of A. Then $D = C_A(B)$ is simple, $B = C_A(D)$, and $[B : k][D : k] = [A : k]$.*

Proposition. *Let A be a CSA over k. Let $L \subset A$ be a (commutative) field containing k. Then TFAE:*

(a) $L = C_A(L)$.

(b) $[A : k] = [L : k]^2$

(c) L *is a maximal commutative subalgebra of* A.

Along with this criterion, the following is an exercise in using the double centralizer theorem.

Corollary (CFT, 3.6). *Let A be a central simple algebra over k. A field L of finite degree over k splits A iff there exists an algebra B representing the same Brauer group element containing L and such that $[B : k] = [L : k]^2$.*

We can give some additional information about splitting fields of A.
If A is any CSA over k with $[A;k] = n^2$, there is a variety/k $\underline{\text{Isom}}(A, M_n)$ which represents the functor

$$\underline{\text{k-algebras}} \to \underline{\text{Sets}} : R \to \{ \text{ isomorphisms } A \otimes_k R \to M_n(R)\}.$$

Indeed, this is easy to represent, since a map of k-algebras is a linear map determined by a non-vanishing determinant and preservation of the algebra's structure constants, all algebraic conditions. The variety is non-empty since it contains at least one $\overline{k}$ point (all CSAs split over an algebraically closed field). Furthermore, it can be checked that this is smooth (by Grothendieck's functorial criterion for smoothness). Hence, the k^{sep} points are dense. In particular, A-splits over some finite separable extension.

CSAs and 2-cocycles

This entire section follows Milne's treatement in **CFT** almost verbatim.

Let L/k be a finite Galois extension.
Let $\mathcal{A}(L/k) = \{A : A = \text{CSA over} k \text{containing} L \text{of degree} [A : k] = [L;k]^2\}$.

By Noether-Skolem, for any $\sigma \in G_{L/K}$, there exists some $e_\sigma \in A^\times$ such that

$$\sigma a = e_\sigma a e_\sigma^{-1} \text{ for all } a \in A. (1)$$

We see that $e_\sigma e_\tau e_{\sigma\tau}^{-1} = \phi_A(\sigma, \tau)$ (1) centralizes L and hence lies in $L^\times$. Because the multiplication in A is associative, we easily see that $\phi_A : G \times G \to L^\times$ is a 2-cocycle. Furthermore, it's clear that different choices of e'_σ lead to a cocycle ϕ'_A which differs from ϕ_A by a coboundary. Thus, we get a well-defined cohomology class.

Claim (CFT, 3.12). *The elements $e_\sigma, \sigma \in G$, are linearly independent over L.*

Proof. $\dim_L(A) = \dim_k(A)/\dim_k(L) = n = |G|$. Thus, it suffices to show that the e_σ are linearly indepedent.
Let $\{e_\sigma\}_J$ be a maximal L-linearly independent subset. We assume, for a contradiction, that $J \neq G$. If $\tau \notin J$, express

$$e_\tau = \sum_{\sigma \in J} a_\sigma e_\sigma. \ (*)$$

for some $a_\sigma \in L$. But we compute $e_\tau a$ in two different ways:
First, by the defining property $e_\tau a e_\tau^{-1} = \tau a$, we have

$$e_\tau a = \tau a e_\tau = \sum_{\sigma \in J} (\tau a) a_\sigma e_\sigma.$$

On the other hand, by our assumption (*) and the definiting property applied to each e_σ, we get

$$e_\tau a = \sum_{\sigma \in J} a_\sigma e_\sigma a = \sum_{\sigma \in J} a_\sigma (\sigma a) e_\sigma.$$

Hence $a_\sigma(\sigma a) = a_\sigma(\tau a)$ for each $\sigma \in J$. But a_σ is non-zero for some $\sigma \in J$, whence $\sigma = \tau$. This contradicts $\tau \notin J$. □

Suppose that A and A' are isomorphic elements of $\mathcal{A}(L/k)$. By Noether-Skolem, we can find an isomorphism $f : A \to A'$ such that $f(L) = L$ and $f|_L$ is the identity map.
Note that if we choose elements $e_\sigma \in A$ which satisfy $(1)_A, (2)_A$ then the elements $f(e_\sigma) \in A'$ satisfy $(1)_{A'}, (2)_{A'}$ with cocycle ϕ_A. Hence, the cohomology class only depends on the isomorphism class of A. Furthermore, if L-bases $\{e_\sigma\} \subset A$ and $\{e'_\sigma\} \subset A'$ both have the same cocycle, then extending $e_\sigma \to e'_\sigma$ by L-linearity clearly gives an isomorphism $A \cong A'$.
Hence, we have an injective map

$$\mathcal{A}(L/k)/\cong \to H^2(L/k).$$

But this map is surjective too. Given a 2-cocycle ϕ, we can just define an algebra by (1) and (2). Namely, let

$$A(\phi) = \oplus_{\sigma \in G} e_\sigma$$

with defining relations

$$e_\sigma a e_\sigma^{-1} = \sigma a \text{ for all } \sigma \in G, a \in L$$

$$e_\sigma e_\tau = \phi(\sigma, \tau) e_{\sigma\tau} \text{ for all } \sigma, \tau \in G.$$

By the cocycle relation for ϕ, it follows that the above defines a k-algebra.

Fact (CFT, 3.13). *$A(\phi)$ is a CSA over k. Furthermore, this construction is a group homomorphism: $[A(\phi)][A(\phi')] = [A(\phi + \phi')]$.*

This isomorphism is actually functorial:

$$\begin{array}{ccc} H^2(L/k) & \xrightarrow{Inf} & H^2(E/k) \\ \downarrow & & \downarrow \\ Br(L/k) & \xrightarrow{\text{inclusion}} & Br(E/k) \end{array}$$

where the vertical maps are $\phi \mapsto A(\phi)$. Since both the Brauer groups (resp. H^2s) are limits under inclusions (resp. inflation maps) of finite Galois extensions L/k, the above diagram implies that there is a canonical isomorphism

$$H^2(k) \xrightarrow{\sim} Br(k).$$

This implies the otherwise unobvious fact that

Corollary. *For any field $k, Br(k)$ is torsion. For any finite extension $L/k, Br(L/k)$ is killed by $[L : k]$.*

Proof. The same results are true for cohomology groups. □

Brauer Groups of Local Fields

Usually, the invariant map of local class field theory is constructed by pure group cohomology. We give an alternate presentation based more directly related to CSAs.

Let D be a central division algebra over non-archimedean local field K, say $n^2 = [D : k]$. Let K have ring of integers $\mathcal{O}_K$, maximal ideal $\mathfrak{p} = (\pi)$, and residue field k of size q.
It has a valuation satisfying the usual properties:

- $|\alpha| = 0$ iff $\alpha = 0$.
- For all $\alpha, \beta \in D, |\alpha\beta| = |\alpha||\beta|$.
- For all $\alpha, \beta \in D, |\alpha, \beta| \leq \max\{|\alpha|, |\beta|\}$.

We define $|\alpha|$ as the scaling effect of right multiplication by α. This is equivalently the absolute value (in K) of the determinant of right multiplication by x as a map from D, as a K-vector space, to itself. Then it is clear that the first and second properties from the above list hold. But the triangle inequality is less obvious.
But this actually reduces to the commutative case.
Indeed, we want to show that if $|x| \leq 1$, then $|1 + x| \leq 1$. But the way we've defined it,

$$|x| = |N_{K[x]/K}(x)|_K^{[D:K[x]]}.$$

So we just need to show that if $|N_{K[x]/K}(x)|_K \leq 1$, then $|N_{K[x]/K}(1 + x)|_K \leq 1$. But this is a result that we know to be true of commutative field extensions. Hence, it's true here too.

Define $|\alpha| = (1/q)^{ord(\alpha)}$, giving the normalized valuation on D.
By the way we've defined the valuation, ord extends the usual valuation on any field extension L/K.
We know that any element $x \in D$ is contained in a field extension $K[x]/K$ of degree $\leq n$ (since any maximal subfield of D has degree n over K). Hence,

$$ord(D^\times) \subset n^{-1}\mathbb{Z}.$$

As usual, we define

$$\mathcal{O}_D = \{\alpha \in D : |\alpha| \leq 1\}$$

$$\mathcal{P} = \{\alpha \in D : |\alpha| < 1\}.$$

The absolute value is discrete and multiplicative. So, just as in the case of fields, any element π of largest absolute value generates the two sided ideal $\mathcal{P}$. And any element of $\mathcal{O}_D$ can be expressed uniquely as $u \times \pi^m$ for some $m \geq 0$. Thus, any two-sided ideal can be expressed uniquely as $\mathcal{P}^m$.
In particular, if $\mathfrak{p}$ denotes the prime ideal of K, then $\mathfrak{p}\mathcal{O}_D = \mathcal{P}^e\mathcal{O}_D$ for some integer e, the ramification index. In particular, $ord(D^\times) = e^{-1}\mathbb{Z}$, implying that $e \leq n$.
Also, if $|\alpha| = 1$ for some $\alpha \in D$, then $\alpha \in \mathcal{O}_D^\times$. Hence, $j = \mathcal{O}_D/\mathcal{P}$ is a finite division algebra, and hence a field. Let $f = [j : k]$. If $j = k[a]$ and α is a lift of a to $\mathcal{O}_D$, then

$$f = [j : k] \leq [K[\alpha] : K] \leq n.$$

Exactly as in the case of commutative fields, we see that $n^2 = ef$. Note that $\mathcal{O}_D$ is a free $\mathcal{O}_K$-module of some rank, say m.

- $\mathcal{O}_D \otimes_{\mathcal{O}_K} K = D$, so $m = n^2$.
- $\mathcal{O}_D \otimes_{\mathcal{O}_K} k = \mathcal{O}_D/\mathfrak{p}\mathcal{O}_D$. Thus, $n^2 = \dim_k \mathcal{O}_D/\mathfrak{p}\mathcal{O}_D$. But we can filter

$$\mathcal{O}_D \supset \mathcal{P} \supset ... \supset \mathcal{P}^e = \mathfrak{p}\mathcal{O}_D.$$

  Each successive quotient, of which there are e, has dimension f as a k-vector space. Hence, $n^2 = ef$.

But since $e, f \leq n$ the equality $ef = n^2$ implies that $e = f = n$.

$$n = [j : k] \leq [K[a] : K]$$

Also, we know that the field $k[a] = j$, so $K[\alpha]$ is an extension of K with residue field j.
But the maximal commutative subfield of D has degree n over K. Thus,

$$n \geq [K[a] : K] \geq [j : k] = n.$$

Thus $K[\alpha]/K$ has both degree and residue degree n. Thus, $K[\alpha]/K$ is unramified.
Since every CSA is in the same class as some division algebra, we know that every CSA is split by an unramified extension. Hence,

$$Br(K) = Br(K^{un}/K).$$

The Local Invariant Map

We can use this Brauer group perspective to directly define the invariant map

$$inv_K : Br(K) \to \mathbb{Q}/\mathbb{Z}$$

and the fundamental class of class field theory.

Any CSA over K is split by some unramified field extension $A \subset L/K$. By Noether Skolem, there is some $\alpha \in A^\times$ such that

$$Frob(x) = \alpha x \alpha^{-1} \text{ for all } x \in G_{L/K}.$$

We define

$$inv_K(A) = ord(\alpha) \ (mod\ \mathbb{Z}).$$

But Frobenii are compatible: if $L' \subset L \subset K$ is a tower of unramified field extensions, then $Frob_{L'/K}|_L = Frob_{L/K}$. Thus, this map does not depend on choice of splitting field.
Also, if A/K is split by L and A'/K is split by L', then $A \otimes_k A'/K$ is split by LL'.
Furthermore, if $Frob(x) = axa^{-1}, Frob(x') = a'xa$, then

$$Frob(x \otimes_k x') = (a \otimes_k 1)(1 \otimes_k a)(x \otimes_k x')(1 \otimes_k a)^{-1}(a \otimes_k 1)^{-1}.$$

Thus, we get a homomorphism from CSAs over K to $\mathbb{Q}/\mathbb{Z}$. Furthermore, $M_n(K) \mapsto 0$, because it is already split over K. Thus,

$$Br(K) \to \mathbb{Q}/\mathbb{Z} : A \mapsto inv_K(A)$$

is a well-defined group homomorphism. The above work we've done also easily shows that

$$Br(L/K) \to 1/[L:K]\mathbb{Z}/\mathbb{Z}$$

for an unramified extension L/K.
Here's the most important example of this:

- Let L/K be an unramified extension of degree n with $\sigma = Frob$. Let ϕ be the 2-cocycle

$$\phi(\sigma^i, \sigma^j) = \begin{cases} 1 & \text{if } i+j \leq n-1 \\ \pi & \text{otherwise.} \end{cases}$$

  This is the 2-cocycle of the fundamental class $u_{L/K} \in H^2(L/K)$. In particular, it maps to $1/n \in 1/n\mathbb{Z}/\mathbb{Z}$ via the invariant map, $inv'_{L/K}$, of Galois cohomology.
  It has associated CSA $A(\phi) = \oplus_i e_i L$ with mutiplication determined by

$$e_i a e_i^{-1} = \sigma^i a \text{ for all } a \in L$$

  and

$$e_i e_j = \begin{cases} e_{i+j} & \text{if } i+j \leq n-1 \\ \pi e_{i+j-n} & \text{otherwise.} \end{cases}$$

  So in particular, e_0 is the identity and L is identified with Le_0. But $e_1^n = e_{n-1}e_1 = \pi e_0 = \pi$. Hence,

$$inv_K(A(\phi)) = ord(e_1) = \frac{1}{n}.$$

Hence, we have a commutative diagram

$$\begin{array}{ccc} H^2(L/K) & \xrightarrow{inv'_{L/K}} & 1/n\mathbb{Z}/\mathbb{Z} \\ \downarrow & & \downarrow \\ Br(L/K) & \xrightarrow{inv_{L/K}} & 1/n\mathbb{Z}/\mathbb{Z} \end{array}.$$

It commutes because it commutes on a generator ϕ of $H^2(L/K)$. Since the top row is an isomorphism, so is the bottom row. Hence, we get a canonical isomorphism

$$inv_K : Br(K) \to \mathbb{Q}/\mathbb{Z}$$

for any local field K.

Brauer Groups of Global Fields

The following is the fundamental exact sequence of class field theory

$$0 \to Br(K) \xrightarrow{\sum_v inv_v} \oplus_v Br(K_v) \to \mathbb{Q}/\mathbb{Z} \to 0.$$

None of the exactness comes easily. In fact, it's not even immediately clear why the first map is well-defined, i.e. why is it impossible for infinintely many of the local invariants to be non-zero simultaneously?
For this, we could return to the variety of ismorphisms $\underline{\mathrm{Isom}}(A, M_n)$. This is a variety over k. If we spread it out to some ring of S-integers $\mathcal{O}_S$, then we'd get a variety $V/\mathcal{O}_S$ which represents the functor

$$\underline{\mathcal{O}_S\text{-algebras}} \to \underline{\mathrm{Sets}} : R \to \{ \text{ isomorphisms } A \otimes_k R \to M_n(R)\}.$$

We could check that this is a non-empty smooth variety. Furthermore, over the residue fields $k_v, v \notin S$ there is a point. By Hensel's Lemma, these lift to $\mathcal{O}_v$ points. (This is a "reason", but not a proof.)

Exactness of the above sequence is the essence of the proofs behind class field theory. For details, see **CFT**.
Another miracle happens over global fields (containing a primitive n^{th} root of 1, call it ζ).
We define the **Milnor K-group**, $K_2(K)$ to be $K^\times \otimes_{\mathbb{Z}} K^\times$ modulo the relation

$$u \otimes_{\mathbb{Z}} (1-u) = 1 \text{ whenever } u, 1-u \neq 1.$$

Now, consider the algebra $A(a, b; \zeta)$ over K generated by i, j subject to the relations:

$$i^n = a, j^n = b, ij = \zeta ji.$$

The Milnor relations are satisfied by the $A(a, b; \zeta)$ and so define a homomorphism

$$K_2(k) \to Br(k).$$

It turns out that

$$K_2(K)/nK_2(K) \to Br(K)[n]$$

is an isomorphism! (Merkuryev-Suslin)
In particular, combining the class field theory exact sequence with this miracle, we see that any 2-torsion element of $Br(K)$ is in the class of a unique quaternion algebra (up to isomorphism).

References

CFT J.S. Milne. *Class Field Theory (version 4.00).* Available at www.jmilne.org/math/

Warning – these notes were written for AV's personal use and have not been checked in any way whatsoever, nor have they been edited for coherence beyond adding one or two sentences. AV makes absolutely no warranty of correctness and these should be used with extreme caution. The talk sketched a proof of Jacquet-Langlands between the quaternion algebra ramified at $\{\infty, 11\}$ and GL_2 at level 11.

1. The adelic quotient associated to a quaternion algebra

D the quaternion algebra ramified at $\{11, \infty\}$. This is represented by $x + yi + zj + wk$, where $i^2 = -11, j^2 = -11, k^2 = -1$; also $ij = 11k, jk = i, ki = j$.

We want to have some "concrete" understanding of the adelic quotient $\mathbf{Q}^\times \backslash (D \otimes \mathbb{A})^\times / \mathbb{A}^\times U$, where U is a maximal compact subgroup of $(D \otimes \mathbb{A})^\times$. It comes with Hecke operators T_p for each prime p.

The maps[1] $q \mapsto \mathrm{Cliff}^0(q), E \mapsto \mathrm{End}(E)$ give bijections:

- Maximal orders in D, i.e. quaternion rings[2] of discriminant -11.
- Isomorphism classes of supersingular elliptic curves over $\overline{\mathbf{F}_{11}}$;
- (Definite) ternary quadratic forms of $\frac{1}{2}$-discriminant -11.

This set is equipped with the structure of a $p+1$-valent directed graph from the Hecke operators; they are represented by matrices

$$A_2 = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix},\ A_3 = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix},\ A_5 = \begin{pmatrix} 4 & 3 \\ 2 & 3 \end{pmatrix},\ A_7 = \begin{pmatrix} 4 & 6 \\ 4 & 2 \end{pmatrix}.$$

Exercise. Why are these matrices not symmetric?

We describe each of the three realizations in turn.

1.1. *Maximal orders.–* Write $\mathfrak{o} = \mathbf{Z}[\frac{1+j}{2}]$. Then $\mathfrak{o}[k]$ is a maximal order.

$$\mathcal{O}_1 := \frac{x + yi + zj + wk}{2} \quad x \equiv z, y \equiv w (2).$$

The group of units (elements of norm 1) is simply $\{\pm 1, \pm k\}$.

Write $\omega = \frac{2+i+k}{4}$, a cube root of 1, and $\nu = \frac{1+i-j-k}{2}$, with norm 6. Then

$$\mathcal{O}_2 := \mathbf{Z} + \mathbf{Z}j + \mathbf{Z}\omega + Z\nu$$

is another maximal order. The group of units is $\{\pm 1, \pm\omega, \pm\omega^2\}$.

$\mathcal{O}_1$ and $\mathcal{O}_2$ intersect in an order that has index 2 in both. We can describe the passage from $\mathcal{O}_1$ to $\mathcal{O}_2$ as follows (with $q = 2\omega$):

$$\mathcal{O}_2 = \frac{q}{2} + \langle z \in \mathcal{O} : \mathrm{tr}(zq) \text{ even} \rangle.$$

The element q has the property that $\mathrm{tr}(q) = 0, \mathrm{N}(q) \equiv 0$ modulo 2.

More generally: Given any maximal order B, there exist $p+1$ neighbours B'. Namely, there exist exactly $p+1$ elements of B/pB of trace and norm 0. For each such x, let $x^* \in B$ be a lift that has norm divisible by p^2. Set $B' = \frac{x^*}{p} + (x^*)^\perp$.

---

[1] Given a ternary quadratic form q the *even Clifford algebra* gives a ternary quadratic form of the same discriminant. For instance

$$ax^2 + by^2 + cz^2 \mapsto \langle i, j, k | i^2 = -bc, j^2 = -ac, k^2 = -ab \rangle$$

Note that upon composition with "trace form" we get the map "multiplication by disc," at least over a field.

[2] A quaternion ring over R is free of rank 4 together with an involution $x \mapsto x^*$ so that the characteristic polynomial is as expected (roots x, x^* with multiplicty 2).

1.2. *Supersingular elliptic curves.*–

The p-neighbour operation is described thus: For each supersingular elliptic curve and prime $p \neq 11$, we can consider the $p+1$ different curves formed by quotienting by a subgroup of order p. This gives a $\delta \times \delta$ matrix whose columns add up to $p+1$, a "Brandt matrix."

There are two supersingular j-invariants in characteristic 11, namely, 1728 and 0. Over $\mathbf{C}$ these are represented by points i, ω. The j-invariant of all points $3i, i/3, (i+1)/3, (i+2)/3$ satisfy the equation $x^2 - 153542016x - 1790957481984 = 0$, which has roots $\{0, 1728\}$ in $\mathbf{F}_{11}$. On the other hand, $j(3\omega) = -12288000$, as are two of the other Hecke translates; the other Hecke translate $(\omega+1)/3$ has j-invariant zero.

There is a 2-isogeny from $y^2 = x^3 - x$ to $y^2 = x^3 + 1$ in characteristic 11.

$$x \mapsto \frac{6x^2 + 5x + 1}{x - 1}, y \mapsto y\frac{x^2 + 9x + 10}{(x-1)^2}.$$

1.3. *Quadratic forms.* – There are two quadratic forms, represented by $q_1 := x^2 + y^2 + 3z^2 - xz, x^2 + y^2 + xy - yz - zx + 4z^2$, and they have respectively 8 and 12 automorphisms. The *total mass* of the genus is $5/24$.

Example of q_1. The associated quaternion algebra spanned by $1, i = e_2e_3, j = e_3e_1, k = e_1e_2$ has relations

$$i^2 = -3, j^2 = -j - 3, k^2 = -1, jk = i^*, ki = j^*, ij = 3k^*$$

For instance $e_3e_1e_3e_1 = e_3^2e_1^2 + 2e_3^2\langle e_1, e_3\rangle = -3 - e_3e_1$. Note that we can write this as $(x - z/2)^2 + y^2 + 11z^2/4$; in other terms, in terms of the basis $e_1' = e_1, e_2' = e_2, e_3' = 2e_3 + e_1$ it is $x^2 + y^2 + 11z^2$. Note that $k' := e_1'e_2' = k, j' := e_3'e_1' = 2j + 1, i' := e_2'e_3' = 2i - k$. These satisfy

$$j'^2 = -11, k'^2 = -1, i'^2 = -13 - 2(ik + ki) \cdots = -11, \ldots$$

Note $ki = j^*$, so also $i^*k^* = j$, that is $ik = j$. So $ki + ik = j + j^* = -1$ and so on. In particular, B_q is isomorphic to the suborder of D spanned by $k, (j-1)/2, (i+k)/2$.

2. The cusp form of weight 11

There is precisely weight 2 one cusp form of weight 11, viz.

$$q\prod(1-q^n)^2\prod(1-q^{11n})^2 = q - 2q^2 - q^3 + 2q^4 + q^5 + 2q^6 - 2q^7\ldots$$

(Note that it is congruent to Δ modulo 11.) This corresponds to the elliptic curve

$$y^2 + y = x^3 - x^2,$$

which is in fact $X_1(11)$ *not* $X_0(11)$. We denote by $N(p)$ the number of points on this curve modulo p. So $N(2, 3, 5, 7) = 5, 5, 5, 10\ldots$ whereas $a_p = -2, -1, 1, -2$.

Exercise. Why are all the numbers are all divisible by 5?

The trace formula. – By the Lefschetz formula, the number of fixed points of T_p is $2(p + 1 - a_p)$. (So, $10, 10, 10, 20$.)

The contribution of each cusp is 2.

The fixed points of T_p on $X_0(11)$ are parameterized by pairs (E, Λ) together with a cyclic p-isogeny $\psi : E \to E$ that fixes Λ. In other terms, E has CM by some order $\mathfrak{o}$ that has an element of norm p. Once we fix $\mathfrak{o}$ and an element $t \in \mathfrak{o}$ of norm p, the remaining question is whether t acting on $\mathfrak{o}/11$ fixes a line; if we suppose that

the $p \neq 11$ and that 11 is unramified in $\mathfrak{o}$, this will be so just when $\mathfrak{o}$ is split at 11; if so, there are two such points.

The orders of small discriminant split at 11 are $\mathbf{Z}[\sqrt{-2}]$ and $\mathbf{Z}[\sqrt{-7}]$. These orders are both of class number 1.

In all cases each solution (up to sign) contributes +2 if unramified at 11, +1 else.

Example. There is an element of norm 2 in $\mathbf{Z}[\sqrt{-2}]$, namely $\pm\sqrt{-2}$; also $\pm\frac{1+\sqrt{-7}}{2}$.

Example. Norm 3: $\pm 1 \pm \sqrt{-2}, \frac{1\pm\sqrt{-11}}{2}$.

Example. Norm 5. We need to solve $20 = x^2+dy^2$; the solutions are $\frac{\pm 1\pm\sqrt{-19}}{2}, \frac{\pm 3\pm\sqrt{-11}}{2}$.

Example. Norm 7. We need to solve $28 = x^2 + dy^2$. The solutions are $\pm 1 \pm \sqrt{-6}, \frac{\pm 3\pm\sqrt{-19}}{2}, \sqrt{-7}$. But $1 + \sqrt{-6}$ contributes +4 (class number two) and $\sqrt{-7}$ also does (two orders). Total $8 + 4 + 4 = 16$.

Example. Norm 13. $\frac{\pm 1\pm\sqrt{-51}}{2}$, same for 43. Class numbers are 2 and 1. Total +12. Finally $\sqrt{-13}$ gives +4 – class number two. total 20. Correct ($a_{13} = 4$).

3. The trace of Brandt matrices

The trace of T_p on the split side is a summation

$$-\sum_{\mathfrak{o},\lambda} h(\mathfrak{o})(1 + \left(\frac{-11}{d}\right)).$$

I'll sketch why the trace of T_p on the quaternionic side is

$$p + 1 + \sum_{\mathfrak{o},\lambda} h(\mathfrak{o})(1 - \left(\frac{-11}{d}\right)). \tag{1}$$

Why are these equal? We need to check

$$\sum_{\mathfrak{o},\lambda} h(\mathfrak{o}) = p + 1,$$

and there are two ways to proceed:

(1) Compute the trace of T_p on *forms of weight* 2*, level* 1.
(2) Use the fact that $(\sum q^{n^2})^3 \sum q^{n^2} = (\sum q^{n^2})^4$.

Sketch of proof of (1)*:* I'll explain it in terms of CM elliptic curves: Suppose $\mathfrak{o}$ is inert or ramified at 11 and $\lambda \in \mathfrak{o}$ has norm p. (Let H be the Hilbert class field, and choose a prime above p). For each ideal class J of $\mathfrak{o}$, the elliptic curve E_J, when reduced mod p, comes equipped with a p-isogeny $E_J \to E_J$, i.e., a loop in the adjacency graph.

Example. Let us reconsider norm 3 from this perspective. It is the norm of $\sqrt{-3}, \frac{1\pm\sqrt{-11}}{2}$ in inert orders.

First, $\sqrt{-3}$. It gives rise to the CM elliptic curve with j-invariant 0, $y^2 = x^3 - 1$. However, there are two CM-maps over $\overline{\mathbb{F}_{11}}$, namely $x \mapsto \zeta x$ and $x \mapsto -\zeta x$.

On the other hand, $\sqrt{-11}$ gives rise a CM elliptic curve of j-invariant -32768. Explicitly, with $a = 4 \times 24 \times 539$ and $b = 16 * 539^2$, it is $y^2 = x^3 - ax - b$, a curve of j-invariant -2^{15}. This curve is the minimal model (!!) and it has conductor $2^4 3^2 7^2 11^2$. On the other hand, over $\mathbf{Q}(\sqrt{-11})$ it becomes the curve

$y^2 = 4x^3 - 24x - \sqrt{539}$, which *does* have good reduction at the prime above 11; it has j invariant $1728 \in \overline{\mathbf{F}_{11}}$.

Remark. You can construct the curve with conductor 121 by twisting. It is

$$y^2 = x^3 - 9504x + 365904.$$

Alternate. If you are proficient with ternary quadratic forms [...]

The file contained a section concerning computations at level 121 (esp. local representations at 11). These have been commented out.

Automorphy, Potentially Automorphy and Langlands Base Change

Alexander Paulin

February 18, 2010

Automorphic Galois Representations

Notation

Let F be a number field (not necessarily totally real) and $n \in \mathbb{N}$. Let S_f, S_∞ denote the set of all finite and infinite places of F respectively. For $v \in S_f \cup S_\infty$ we denote the completion by F_v and for $v \in S_f$ we denote by $\mathcal{O}_v$ the ring of integers. Let $\mathbb{A}_F$ denote the adeles over F and $\mathbb{A}_F^\infty$ their finite component. Fix algebraic closures $\bar{F}$ and $\bar{F}_v$ and let $G_F := Gal(\bar{F}/F)$ and $G_{F_v} := Gal(\bar{F}_v, F_v)$. We fix embeddings $\bar{F} \subset \bar{F}_v$, which induces embeddings $G_{F_v} \subset G_F$. For $S \subset S_f$ a finite subset let $G_{F,S}$ denote the Galois group of the maximal extension of F (contained in $\bar{F}$) unramified outside S. For $v \in S_f \setminus S$ let $frob_v \in G_{F,S}$ denote the frobenius element at v.

Automorphic Representations of $GL_{n/F}$

Let π be an automorphic representation of $GL_{n/F}$. Rather than defining such a representation let us recall its predominant features:

1. π is a complex vector space with a tensor decomposition

$$\pi = \pi_f \bigotimes (\otimes_{v \in S_\infty} \pi_{\infty,v})$$

where π_f is an irreducible, smooth, admissible representation of $GL_n(\mathbb{A}_F^\infty)$ and $\pi_{\infty,v}$ is an irreducible, admissible Harish-Chandra module associated to the Lie group $GL_n(F_v)$.

2. We have the restricted tensor decomposition

$$\pi_f = \otimes'_{v \in S_f} \pi_{f,v}$$

where $\pi_{f,v}$ is an irreducible, smooth, admissible representation of $GL_n(F_v)$ $(v \in S_f)$. Moreover, for almost all finite places $\pi_{f,v}$ is unramified, i.e. posesses a $GL_n(\mathcal{O}_{F_v})$ fixed vector. Let $S_\pi \subset S_f$ denote the set of finite places of F for which $\pi_{f,v}$ is ramified.

Recall that for any $v \in S_f$ the Satake Isomorphism induces a natural bijection:

$$\{ \text{ Semisimple conjugacy classes of } GL_n(\mathbb{C}) \ \}$$

$$\updownarrow$$

$$\{ \text{ Isomorphism classes of unramified, irreducible, smooth admissible complex representations of } GL_n(F_v). \ \}$$

Hence for every $v \notin S_\pi$ we get the semisimple conjugacy class $\Upsilon_v \subset GL_n(\mathbb{C})$. By Strong Multiplicity One the set $\{\Upsilon_v\}_{v \notin S_\pi}$ determines π up to isomorphism.

Fix $k \in \mathbb{Z}$. In the case where $n = 2$ and F is totally real if we demand that the Harish-Chandra modules at ∞ are all holomorphic discrete series of weight k, then π is associated to a cuspidal Hilbert modular newform g of parallel weight $k \geq 2$. If $v \notin S_\pi$ and $\mathbb{T}_v$ is the usual Hecke operator at v then if $\mathbb{T}_v g = a_v g$ we have the equality

$$a_v = trace(\Upsilon_v).$$

Now fix a prime $p \in \mathbb{N}$ and an algebraic closure $\bar{\mathbb{Q}}_p$. Also fix an isomorphism $\varsigma : \bar{\mathbb{Q}}_p \cong \mathbb{C}$. Let ρ be a continuous (for the p-adic topology) representation

$$\rho : G_F \to GL_n(\bar{\mathbb{Q}}_p),$$

which is unramified outside the finite set $S_\rho \subset S_f$. Hence ρ factors through G_{F,S_ρ}. For $v \in S_f \setminus S_\rho$ let $\phi_v \subset GL_n(\mathbb{C})$ denote the semisimple conjugacy class associated to $\rho(Frob_v)$. Here was are implicitly using the isomorphism ς. The information $\{\phi_v\}_{v \notin S_\rho}$ determines ρ up to isomorphism by Tchebotarev density.

Definition. *ρ is automorphic if there exists π, an automorphic representation of $GL_{n/F}$ and a finite subset $S \subset S_f$ such that*

1. *$S_\pi \subset S$ and $S_\rho \subset S$.*
2. *$\{\phi_v\}_{v \notin S} = \{\Upsilon_v\}_{v \notin S}$.*

We say that ρ is potentially automorphic if there exists a finite extension E/F, contained in $\bar{F}$, such that $\rho|_{G_E}$ is automorphic.

Global Langlands Reciprocity Conjecture. *If ρ is semisimple, deRham (at all places over p) and unramified outside of a finite set of primes then ρ is automorphic.*

Of course conjecturally all local data should match under the Local Langlands correspondence.

In the Hilbert case this implies that the traces of Frobenius away from S_ρ arise as eigenvalues of Hecke operators on of the space of Hilbert modular forms.

Now let F be totally real and D a quaterion algebra over F. If ρ is 2-dimensional then we say it is automorphic for D if we can find π_D, an automorphic representation of $D_{/F}$, such that ρ is associated to π_D^{JL}, its Jacquet-Langlands transfer. Informally this is saying that the traces of Frobenius away from S_ρ arise as eigenvalues of Hecke operators of the space of automorphic forms associated to D. Of course there is a converse to this - to any cuspidal Hilbert eigenform π of parallel weight $k \geq 2$ there is an associated 2-dimensional p-adic representation of G_F.

1 Langlands Base Change

Let E/F be a finite extension (contained in $\bar{F}$). We have the natural inclusion $G_E \subset G_F$. If ρ is a representation as in the Global Reciprocity Conjecture then $\rho|_{G_E}$ is also of this form. If we where to believe the reciprocity conjecture then if π is associated to ρ there should be an automorphic representation of $GL_{n/E}$, denoted π', associated to $\rho|_{G_E}$.

This "base change" transfer is also predicted by Langlands principle of Functoriality. More precisely, let $G = Res_{E/F}(GL_{n/E})$. The L-group of this groups is

$$^LG = (\prod_{G_E \backslash G_F} GL_n(\mathbb{C})) \rtimes G_F$$

where the G_F acts by permutations. Thus there is a natural L-homomorphism

$$^LGL_{n/F} = GL_n(\mathbb{C}) \times G_F \rightarrow^L G$$

which is the diagonal embedding. Functoriality in this case would transfer an automorphic representation of $GL_{n/F}$ to an automorphic representation of G. However, observing that $G(\mathbb{A}_F) = GL_n(\mathbb{A}_E)$, we see that an automorphic representation of this latter group is simply an automorphic representation of $GL_{n/E}$.

We have the following:

1. For E/F a solvable extension functoriality has been established in this case by Langlands ($n = 2$) and by Aurthur and Clozel ($n > 2$). They use the trace formula to establish functoriality for cyclic extensions of prime degree, then the result follows by induction. We denote the functorial transfer by $BC_{E/F}$ (BC for base change).

2. If E/F is solvable and π is an automorphic representation of $GL_{n/F}$ associated to a Galois representation ρ then $\rho|_{G_E}$ is automorphic with associated representation $BC_{E/F}(\pi)$.

In the Hilbert case we have the converse:

Theorem. *Let E/F be a solvable extension of totally real number fields. Let ρ be a 2-dimensional p-adic representation of G_F. Suppose that $\rho|_{G_E}$ is automorphic with associated representation π', a cuspidal automorphic representation of $GL_{2/E}$ all of who infinite components are holomorphic discrete series of weight $k \geq 2$ (i.e. coming from a Hilbert eigenform over E of parallel weight k). Then there is a cuspidal automorphic representation of $GL_{2/F}$ all of whose infinite components are discrete series of weight k such that $\pi' = BC_{E/F}(\pi)$ (i.e. coming from a Hilbert eigenform over F of parallel weight k).*

The proof of this result relies on a determining the image of the base change map and then determining the fibres. Because π' is associated to a representation of G_E which extends to G_F it is invariant under $Gal(E/F)$. This ensures that it is the base change of some automorphic representation of $GL_{n/F}$. The trick then is to find a lift with the desired properties.

Corollary. *If ρ is a two dimensional p-adic representation of G_F what becomes automorphic (of parallel weight $k \geq 2$) over a solvable totally real extension then ρ is already automorphic (of parallel weight k).*

Hence constructing solvable extensions will be important to us. In particular it will be important to construct solvable global extensions with prescribed local behaviour. In this regard we have the following theorem (an extension of the classical Grunwald-Wang theorem):

Theorem. *Let F be a totally real number field and $\{v_1, \cdots, v_n\} \subset S_f$. Let K_i/F_{v_i} be a finite Galois extension for each i. Then there exists a totally real solvable extension E/F including K_i/F_{v_i} as the local extension at v_i for each i, such that E/F can be chosen to be unramified at any auxilliary finite set of places.*

For a proof of this see Brians number theory handout.

Lecture 16: Review of representation theory

Andrew Snowden
February 26, 2010

In the first (and main) part of these notes, I review the representation theory we have done this semester, highlighting the points that are of most relevance to us. Then I will state a modularity lifting theorem and make a few remarks about how the rerpesentation theory is used in the proof. In my next talk, I will give an outline of the proof of this modularity lifting theorem.

1. Representation theory

The lectures we have had on representation theory centered around these topics:

- The theory of admissible representations of $\mathrm{GL}(2, \mathbf{Q}_p)$ (or more generally, $\mathrm{GL}(2, F)$ with $F/\mathbf{Q}_p$ a finite extension).
- The theory of automorphic representations of GL(2); in particular, the correspondence between Hecke eigenforms in the classical sense and automorphic representations.
- The Jacquet-Langlands correspondence, relating automorphic forms on GL(2) with those on a division algebra.
- Base change, relating automorphic forms on GL(2) over two different fields (one a solvable extension of the other).

I will go through each of these four topics and remind us of the key points for our applications. I will also throw in some material abouth the Langlands correspondence (both local and global) that we may not have covered.

1.1. **Admissible representations.** Let $F/\mathbf{Q}_p$ be a finite extension and let G be the group $\mathrm{GL}(2, F)$. Fix an algebraically closed field K of characteristic zero (one always takes K to be the complex numbers or the closue of some $\mathbf{Q}_\ell$). A representation of G on a K-vector space V is *smooth* if the stabilizer of any vector in V is an open subgroup of G; it is *admissible* if it is smooth and for every open subgroup U of G the space V^U is finite dimensional. We are most interested in irreducible admissible representations. Here "irreducible" has its usual sense: the only stable subspaces are 0 and the whole space.

An easy way to construct admissible representations is through induction. Let $\alpha, \beta : F^\times \to K^\times$ be two continuous characters. Continuity amounts to the condition that the restriction of α and β to the group of units U_F should factor through a finite quotient of U_F. Let $V = V(\alpha, \beta)$ be the space of all locally constant functions $f : G \to K$ which satisfy the identity

$$f\left(\begin{pmatrix} a & x \\ & b \end{pmatrix} g\right) = \alpha(a)\beta(b)\left|\frac{a}{b}\right|^{1/2} f(g)$$

for all $a, b \in F^\times$, $x \in F$ and $g \in G$. We let G act on V by right translation: $(gf)(g') = f(g'g)$. It is quite easy to see that this makes V into an admissible representation of V. A more difficult result is the following: if $\alpha\beta^{-1}$ is not equal to $|\cdot|$ or $|\cdot|^{-1}$ then V is irreducible. Here $|\cdot|$ is the norm character of $F^\times$, which takes $a \in F^\times$ to $q^{-\operatorname{val} a}$ where q is the cardinality of the residue field. These irreducible admissible representations are called the *principal series*.

When $\alpha\beta^{-1}$ is equal to $|\cdot|^{\pm 1}$ the representation $V(\alpha, \beta)$ is no longer irreducible. Rather, it is indecomposable and has two Jordan-Holder constituents. One of these constituents is one dimensional while the other is infinite dimensional. Precisley, say $\alpha\beta^{-1} = |\cdot|$ and write $\alpha = \gamma|\cdot|^{1/2}$ and $\beta = \gamma^{-1}|\cdot|^{-1/2}$. Then $V(\alpha, \beta)$ has a unique irreducible subrepresentation $\mathrm{St}(\gamma)$ which is infinite dimensional. The quotient $V(\alpha, \beta)/\mathrm{St}(\gamma)$ is one dimensional and G acts on it through the character $g \mapsto \gamma(\det g)$. Write St in place of $\mathrm{St}(\gamma)$ where γ is the trivial character. The representation St is called the *Steinberg representation*. One has $\mathrm{St}(\gamma) = \mathrm{St} \otimes \gamma$.

We have thus completely analyzed the representations $V(\alpha, \beta)$. There are many irreducible admissible representations of G which do not appear inside of these reperesentations, however; these are called the *supercuspidal representations* of G. We now have the following classification of the irreducible representations of G.

Theorem 1.1. *Let V be an irreducible admissible representation of G over K. Then V is equivalent to one and only one of the following:*

- *An irreducible principal series* $V(\alpha, \beta)$ *with* $\alpha\beta^{-1} \neq |\cdot|^{\pm 1}$.
- *A one dimensional representation corresponding to a character* $g \mapsto \gamma(\det g)$.
- *A twist* $\mathrm{St} \otimes \gamma$ *of the Steinberg representation* St.
- *A supercuspidal representation.*

This theorem almost follows by our definition of supercuspidal. The one part that does not is its assertion that the principal series and twists of Steinbergs are inequivalent. The one dimensional representations are often counted as principal series. We will sometimes treat them as such and sometimes not.

An irreducible admissible representation V of G is called *unramified* if it has a vector which is invariant under the maximal compact subgroup $\mathrm{GL}(2, \mathscr{O}_F)$. It is a theorem that V is unramified if and only if it is a principal series of the form $V(\alpha, \beta)$ with α and β unramified characters of $F^\times$ (where here unramified means trivial on U_F), or a one dimensional principal series given by $g \mapsto \gamma(\det g)$ with γ unramified. Note that an unramified character of $F^\times$ is determined by a single number, namely, its value on any uniformizer.

Key points: (1) The irreducible admissible representations of G fall into three classes: principal series, twists of Steinberg and supercuspidal. (2) The unramified representations of G are exactly the principal series representations coming from unramified characters. These are parameterized by (unordered) pairs of numbers (elements of $K^\times$).

1.2. **The local Langlands correspondence.** Keep the notation of the previous section. We have an exact sequence

$$0 \to I_F \to \mathrm{Gal}(\overline{F}/F) \stackrel{\mathrm{val}}{\to} \widehat{\mathbf{Z}} \to 0$$

where I_F is the inertia subgroup of the Galois group. The *Weil group* of F is by definition the subgroup of $\mathrm{Gal}(\overline{F}/F)$ given by $\mathrm{val}^{-1}(\mathbf{Z})$. We call a representation of W_F on a K-vector space V *Frobenius semi-simple* if some fixed Frob in W_F acts semi-simply. Recall that a *Weil-Deligne representation* of F with coefficients in K is a pair (V, N) where:

- V is a K vector space with an action of W_F which is Frobenius semi-simple and under which inertia acts through a finite quotient.
- N is an endomorphism of V which satisfies

  $$gNg^{-1} = q^{\mathrm{val}\, g} N$$

  where q denotes the cardinality of the residue field of F. Equivalently, N defines a W_F-equivariant map $V(1) \to V$ where $V(1)$ is the twist of V by the character $g \mapsto q^{\mathrm{val}\, g}$.

The collection of all Weil-Deligne representations forms a category and this cateogry is abelian. The following theorem is not difficult:

Theorem 1.2. *Let* $\ell \neq p$ *be a prime number. There is then an equivalence of categories:*

$$\left\{ \begin{array}{l} \textit{Weil-Deligne representations} \\ \textit{with coefficients in } \overline{\mathbf{Q}}_\ell \end{array} \right\} \leftrightarrow \left\{ \begin{array}{l} \textit{Continuous Frobenius semi-simple representa-} \\ \textit{tions of } W_F \textit{ on } \overline{\mathbf{Q}}_\ell \textit{ vector spaces} \end{array} \right\}$$

Sketch of proof. Let (V, N) be a Weil-Deligne representation. Let ρ denote the action of W_F on V. Define a new representation ρ' of W_F on V by

$$\rho'(\mathrm{Frob}^n g) = \rho(\mathrm{Frob}^n g) \exp(N t_\ell(g)).$$

Here $\mathrm{Frob} \in W_F$ is a fixed Frobenius element, g is an element of the inertia subgroup I_F of W_F and $t_\ell : I_F \to \mathbf{Z}_\ell$ is the tame ℓ-adic character. One easily verifies that ρ' is a continuous Frobenius semi-simple representation. We have thus defined a map of categories. One must then check that it is in fact an equivalence, which is not difficult. □

It is not difficult to classify two dimensional Weil-Deligne representations:

Theorem 1.3. *Let* (V, N) *be a two dimensional Weil-Delgine representation of* F *with coefficients in* K. *Then* (V, N) *falls into exactly one of the following three cases:*

- V *is a direct sum of two characters of* W_F *and* $N = 0$.
- V *is irreducible under* W_F *and* $N = 0$.
- V *is a direct sum* $W \oplus W(1)$ *where* W *is one dimensional (and thus acted on by a character* γ *of* W_F*);* N *kills* $W(1)$ *and maps* W *isomorphically onto* $W(1)$.

We can now state a version of the local Langlands correspondence for GL(2).

Theorem 1.4. *There is a natural bijection*

$$\left\{\begin{array}{l}\text{Irreducible admissible represen-}\\ \text{tations of } \mathrm{GL}(2,F) \text{ over } K\end{array}\right\} \leftrightarrow \left\{\begin{array}{l}\text{Two dimensional Weil-Deligne repre-}\\ \text{sentations with coefficients in } K\end{array}\right\}.$$

Under thie bijection, the principal series correspond to direct sums of characters, the supercuspidals to irreducibles and the twists of Steinberg to the Weil-Deligne representations with non-zero N. More precisely, the principal series $V(\alpha,\beta)$ corresponds to the representation $\alpha' \oplus \beta'$ where α' and β' correspond to α and β by class field theory. One can make a similar statement for twists of Steinberg.

Key points: (1) Two dimensional Weil-Deligne representations fall into three classes. (2) There is a natural bijection between two dimensional Weil-Deligne representations and irreducible admissible representations of $\mathrm{GL}(2,F)$. This bijection preserves the trichotomy on each side and on principal series and twists of Steinberg can be computed in terms of class field theory. (3) Weil-Deligne representations basically correspond to continuous ℓ-adic representations of the Weil group for any $\ell \neq p$, and these are almost the same thing as representations of the absolute Galois group.

1.3. **Automorphic representations.** Now let F be a number field and let $\mathbf{A}_F$ be its adele ring. An *automorphic form* on $\mathrm{GL}(2)$ over F is a function $f : \mathrm{GL}(2,\mathbf{A}_F) \to \mathbf{C}$ satisfying a number of properties, the most important of which is that it is invariant on the left under $\mathrm{GL}(2,F)$. The set of all automorphic forms forms a vector space $\mathscr{A}_F$. This vector space carries an action of $\mathrm{GL}(2,\mathbf{A}_F^f)$ by right translation. Furthermore, the Lie algebra and the maximal compact of $\mathrm{GL}(2,F_\infty)$ act on $\mathscr{A}_F$ (that is, $\mathscr{A}_F$ is a Harish-Chandra module for $\mathrm{GL}(2,F_\infty)$). (The full group $\mathrm{GL}(2,F_\infty)$ does not act on $\mathscr{A}_F$ as it destroys the K-finiteness condition.) An *automorphic representation* of $\mathrm{GL}(2,\mathbf{A}_F)$ is something of the form $\pi_f \otimes \pi_\infty$ where π_f is an irreducible admissible representation of $\mathrm{GL}(2,\mathbf{A}_F^f)$ and π_∞ is an irreducible Harish-Chandra module of $\mathrm{GL}(2,F_\infty)$ such that $\pi_f \otimes \pi_\infty$ is equivalent to a submodule of $\mathcal{A}_F$. There is a certain condition called *cuspidal* that one can impose on automorphic forms. The set of all cuspidal forms forms a vector subspace $\mathscr{A}_F^\circ$ of $\mathscr{A}_F$ which is stable under the various actions of pieces of $\mathrm{GL}(2,\mathbf{A}_F)$. An automorphic representation is *cuspidal* if it appears inside this cuspidal space.

Say for the moment that $F = \mathbf{Q}$. As we have discussed earlier in the semester, classical modular eigenforms correspond bijectively to automorphic representations π for which π_∞ is a discrete series representation. More precisely, say f is a newform of level N and weight k and let π be the corresponding automorphic representation. We can then write $\pi = \pi_f \otimes \pi_\infty$ and further decompose π_f as a restricted tensor product $\otimes \pi_p$, where π_p is an irreducible admissible representation of $\mathrm{GL}(2,\mathbf{Q}_p)$. The Harish-Chandra module π_∞ is completely determined by the weight k. For primes p not dividing the level, π_p is an unramified representation of $\mathrm{GL}(2,\mathbf{Q}_p)$. As we have seen, such representations are determined by two numbers; the representation π_p corresponds to the eigenvalues of the Hecke operators T_p and $T_{p,p}$ acting of f. (There is a precise formula to take these two numbers and produce two characters α and β of $\mathbf{Q}_p^\times$ such that π_p is equivalent to $V(\alpha,\beta)$.) For primes p dividing N the representation π_p is *not* unramified. I imagine that it is possible to determine π_p from a classical point of view; however, this is probably a bit complicated. This is one of the main advantages of the formulation in terms of automorphic representations: the information at ramified primes is more readily accessible.

When $F \neq \mathbf{Q}$ the discussion of the previous paragraph carries over but is a bit more complicated. The reason that it becomes more complicated is that the corresponding classical picture becomes more complicated. For example, in the setting of Hillbert modular forms the space which plays the role of the modular curve can be disconnected: it will be a disjoint union of spaces of the form $\mathfrak{h}^n/\Gamma_i$ where $\mathfrak{h}$ is the upper half plane and the Γ_i are certain arithmetic groups. The proper analogue of a modular form is then a tuple (f_i) where f_i is a function on $\mathfrak{h}^n$ invariant under Γ_i. The Hecke operators then permute the f_i in addition to acting in the usual fashion. This additional bookkeeping required makes the classical point of view much more cumbersome to deal with. It is another reason for swithcing to the representation theoretic perspective.

Key points: (1) Classical modular forms correspond to automorphic representations of $\mathrm{GL}(2,\mathbf{A_Q})$ satisfying a certain condition at infinity. (2) Automorphic representations are built out of irreducible admissible representations at each finite place and a Harish-Chandra module at infinity. Almost all of these irreducible

admissible representations are unramified and the two parameters that determine them correspond to the two Hecke eigenvalues in the classical picture. (3) Automorphic representations are much better to deal with for certain applications: even in the most basic case of classical modular forms they give easier access to information at ramified primes; in more complicated situations, they remove the cumbersome bookkeeping that is present in the classical picture.

1.4. **The global Langlands correspondence.** Let f be a modular form on the upper half plane of weight k and level N which is an eigenform for the Hecke operators T_p and $T_{p,p}$ away from N. As we discussed in the first semester, there is a Galois representation

$$\rho_{f,\ell} : G_{\mathbf{Q}} \to \mathrm{GL}_2(\overline{\mathbf{Q}}_\ell)$$

which satisfies and is uniquely determined by the following property: if $p \neq \ell$ is a prime not dividing N then $\rho_{f,\ell}$ is unramified at p and the characteristic polynomial of $\rho_{f,\ell}(\mathrm{Frob}_p)$ is given by $T^2 - a_p T + a_{p,p}$ where a_p and $a_{p,p}$ are the eigenvalues of f under T_p and $T_{p,p}$. The representation $\rho_{f,\ell}$ is "odd," that is, its determinant on a complex conjugation is -1.

As we have seen, in certain situations it is better to use automorphic representations in place of modular forms. This is one of those situations! The above result can be generalized and refined, and to state the improved version it is better to use automorphic representations. Let F be a totally real number field and let π be an automorphic representation of $\mathrm{GL}(2, \mathbf{A}_F)$ such that π_∞ is a discrete series representation. Then there is a Galois representation

$$\rho_{\pi,\ell} : G_F \to \mathrm{GL}_2(\overline{\mathbf{Q}}_\ell)$$

which satisfies and is uniquely determined by the following property: if v is a place of F which does not lie above ℓ then $\rho_{\pi,\ell}|G_{F,v}$ corresponds to π_v under the local Langlands correspondence. The representation $\rho_{\pi,\ell}$ is also odd: its determinant on any complex conjugation is -1. (Note that the condition that π_∞ be discrete series is equivalent to the condition that the corresponding classical modular form be holomorphic.)

The above result is clearly more general than the first one since it permits F to be a totally real field rather than just $\mathbf{Q}$. However, even for $F = \mathbf{Q}$ it is a stronger result: it specified the local Galois representation everywhere except at ℓ in terms of the corresponding local component of the automorphic representation. The local Galois representation at ℓ is much more subtle: it is not determined by the corresponding component of the automorphic representation.

It is expected that the $\rho_{\pi,\ell}$ give all the Galois representations which are odd, ramified at finitely many places and satisfy some local condition at ℓ (coming from ℓ-adic Hodge theory). This has basically been proved for $F = \mathbf{Q}$ but is still open for all other F. The most critical intermediate result in the proof for $F = \mathbf{Q}$ is a modular lifting theorem; we will prove such a theorem in this seminar.

Key point: Given an automorphic representation π of a toally real number field which is discrete series at infinity, there is a corresponding Galois representation $\rho_{\pi,\ell}$. (Or rather, one for each ℓ.) The restriction of $\rho_{\pi,\ell}$ to a decomposition group away from ℓ corresponds to the local component of π under the local Langlands correspondence. Furthermore, $\rho_{\pi,\ell}$ is an odd representation.

1.5. **The Jacquet-Langlands correspondence.** Let F be a number field. Let G be the algebraic group $\mathrm{GL}(2)$ over F. Let D be a quaternion algebra over F and let G' be its unit group, regarded as an algebraic group (so $G'(A) = (D \otimes_F A)^\times$). One then has the notion of an automorphic representation of G'. The global Jacquet-Langlands correspondence is the following theorem:

Theorem 1.5. *The is a natural bijection:*

$$\{\textit{Automorphic representations of } G'\} \leftrightarrow \left\{ \begin{array}{l} \textit{Automorphic representations of } G \\ \textit{which are essentially square integrable} \\ \textit{at all places where } D \textit{ ramifies} \end{array} \right\}$$

(An irreducible admissible representation of $\mathrm{GL}(2, F_v)$ *is* essentially square integrable *if it is a twist of the Steinberg or supercuspidal, i.e., not principal series.) Furthermore, if π' is an automorphic representation of G' and π the corresponding automorphic representation of G then π_v is determined completely by π'_v. Two special cases: (1) if D splits at v and we identify D_v with $M_2(F_v)$ then π'_v is identified with π_v; (2) if π'_v is the trivial representation then π_v is the Steinberg representation.*

Assume that D is ramified at all infinite places; this is the case we care most about. For a compact open subgroup U of $(D \otimes \mathbf{A}_F^f)^\times$ let $S_2(U)$ denote the space of all functions

$$D^\times \backslash (D \otimes \mathbf{A}_F^f)^\times / U \to \mathbf{C}.$$

Note that the double quotient above is a finite set; we really do mean *all* possible functions, there is no possible continuity condition to impose. For a place v of F at which U is maximal compact and D is split there is a natural Hecke operator T_v that acts on $S_2(U)$. The Jacquet-Langlands correspondence implies that if f is a parallel weight 2 holomorphic cuspidal Hilbert eigenform whose associated automorphic representation is essentially square integrable at the places where D is ramified then there is an element g of $S_2(U)$ which is an eigenvector for all the Hecke operators and has the same eigenvalues as f. (Here U is determined from the level of f.) Therefore, as long as we are in a situation where the appropriate local conditions are in place, we can work with $S_2(U)$ instead of the space of Hilbert modular forms. This is advantageous because functions on a finite set are very easy to think about! For instance, there is an obvious integral structure on $S_2(U)$ (take integral valued functions) and so the notion of a mod p modular form on D is evident.

Key points: (1) One can move automorphic forms and representations between GL(2) and quaternion algebras; the only obstructions are local and fairly simple. (2) By taking D to be ramified at infinity, automorphic forms on D can be thought of as functions on a finite set.

1.6. **Base change.** Let π be an automorphic representation of $\mathrm{GL}(2, \mathbf{A}_F)$ with F a number field, such that π_∞ is discrete series. As we have seen, there is then an associated Galois representation $\rho_{\pi,\ell}$. Given an extension F'/F we can restrict $\rho_{\pi,\ell}$ to $G_{F'}$. This is the sort of Galois representation that we expect is of the form $\rho_{\pi',\ell}$ for some automorphic representation π' of $\mathrm{GL}(2, \mathbf{A}_{F'})$. The automorphic representation π' has been proven to exist when the extension F'/F is solvable. Precisely we have the following:

Theorem 1.6. *Let F'/F be a solvable extension of number fields. There is a natural map of sets*

$$\mathrm{BC} : \left\{ \begin{matrix} \textit{Automorphic representations} \\ \textit{of } \mathrm{GL}(2, \mathbf{A}_F) \end{matrix} \right\} \to \left\{ \begin{matrix} \textit{Automorphic representations} \\ \textit{of } \mathrm{GL}(2, \mathbf{A}'_F) \end{matrix} \right\}$$

such that if $\pi' = \mathrm{BC}(\pi)$ then: (1) the local component π'_v can be computed in terms of π_v; (2) if π_∞ is discrete series then so is π' and $\rho_{\pi',\ell} = \rho_{\pi,\ell}|_{G_{F'}}$.

There is a local base change map also: if F'_v is a finite extension of F_v then there is a base change map BC from irreducible admissible representations of $\mathrm{GL}(2, F_v)$ to those of $\mathrm{GL}(2, F'_v)$. In fact, the meaning of (1) in the above theorem is precisely that $\pi'_v = \mathrm{BC}(\pi_v)$. Thus local and global base change are compatible. The local base change map satisfies a property analogous to (2) above, namely, it commutes with the local Langlands correspondence.

From the above properties of local base change, and what we know about local Langlands, it is easy to see some examples of how local base change works. For example, the principal series $V(\alpha, \beta)$ corresponds under local Langlands to the Galois representation $\alpha' \oplus \beta'$ where α' and β' correspond to α and β under class field theory. Restricting this to $G_{F'_v}$ we simply get $\alpha'|_{G_{F'_v}} \oplus \beta'|_{G_{F'_v}}$. Going the other way under local Langlands, this corresponds to the principal series $V(\alpha'', \beta'')$ where α'' and β'' correspond to $\alpha'|_{G_{F'_v}}$ and $\beta|_{G_{F'_v}}$ under class field theory. Now, class field theory turns restriction to a larger number field into composition with the norm. Thus $\alpha'' = N^*\alpha$ and $\beta'' = N^*\beta$, where $N : (F')^\times \to F^\times$ is the norm map. We thus find

$$\mathrm{BC}(V(\alpha, \beta)) = V(N^*\alpha, N^*\beta).$$

The base change of a principal series is always a principal series. Similarly, the base change of a twist of Steinberg is again a twist of Steinberg — restricting to a bigger field will never turn a non-zero N zero or vice versa. By this reasoning, the base change of a supercuspidal will never be a twist of Steinberg. However, an irreducible Galois representation can certainly restrict to a reducible one. Thus it is possible for the base change of a supercuspidal to be principal series. In fact, if π is any irreducible admissible representation of $\mathrm{GL}(2, F_v)$ then one can find an extension F'_v/F_v such that $\mathrm{BC}(\pi)$ is either unramified or Steinberg. Any base change of Steinberg is still Steinberg, however.

The above local discussion has the following global application (when compbined with some global class field theory). Given an automorphic representation π of $\mathrm{GL}(2, \mathbf{A}_F)$ there exists a finite solvable Galois

extension F'/F such that the base change of π to F' is everywhere unramified or Steinberg. In fact, if F is totally real (as it will be in our applications) then F' can be taken to be totally real as well.

There is a sort of converse to base change that will be useful for us, which we refer to as *solvable descent*.

Theorem 1.7. *Let F be a totally real number field and let $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ be a Galois representation. Assume that there exists a finite, totally real, solvable extension F'/F and a parallel weight 2 automorphic representation π' of $\mathrm{GL}(2, \mathbf{A}_{F'})$ such that $\rho|_{G_{F'}} = \rho_{\pi',p}$ and both are irreducible. Then there exists a parallel weight 2 automorphic representation π of $\mathrm{GL}(2, \mathbf{A}_F)$ such that $\rho = \rho_{\pi,p}$.*

In other words: if ρ becomes modular over a solvable extension then ρ is modular.

Key points: (1) There is an operation ("base change") on automorphic representations and local representations which corresponds to restriction on the Galois side, at least for solvable extensions. (2) Given an automorphic representation, one can always make a solvable base change such that the result is either unramified or Steinberg at all places. One cannot get rid of Steinbergs through base change, however. (3) Given a Galois representation, one can check if it comes from an automorphic form by checkings over a solvable extension (subject to some technicalities).

2. Modularity lifting

We will now state a modularity lifting theorem that we will later use and indicate how base change and the Jacquet-Langlands correspondence are used in the proof. We must first make some Galois theoretic definitions.

Let $F/\mathbf{Q}_p$ be a finite extension. We say that a Galois representation $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ is *ordinary* if it is of the form

$$\begin{pmatrix} \alpha\chi_p & * \\ & \beta \end{pmatrix}$$

where α and β are finitely ramified characters, and, as always, χ_p denotes the p-adic cycloctomic character. (One could allow for more general definitions of ordinary, replacing χ_p by χ_p^n; for now we will stick with this one.) Let E/F be an extension over which α and β become unramified. The representation $\rho|_{I_E}$ is an extension of the trivial representation by χ_p and so defines an element of $H^1(I_E, \overline{\mathbf{Q}}_p(\chi_p))$, which is identified with $\overline{\mathbf{Q}}_p \otimes (E^{\mathrm{un}})^\times$ by Kummer theory. (Here E^{un} is the maximal unramified extension of E and I_E is the inertia subgroup of G_E.) We say that ρ is *potentially crystalline* if this class belongs to $\overline{\mathbf{Q}}_p \otimes \mathscr{O}_{E^{\mathrm{un}}}^\times$. This is idendependent of the choice of E.

Now let $F/\mathbf{Q}$ be a finite totally real extension. Recall that a representation $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ is *odd* if $\det \rho(c) = -1$ for all complex conjugations $c \in G_F$. We can now state a modular lifting theorem.

Theorem 2.1. *Let $p > 5$. Let $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ be an odd, finitely ramified representation such that $\overline{\rho}|_{G_{F(\zeta_p)}}$ is absolutely irreducible and ρ is potentially crystalline and ordinary at all places above p. Assume that there exists an automorphic representation π of $\mathrm{GL}(2, \mathbf{A}_F)$ such that $\rho_{\pi,p}$ is potentially crystalline and ordinary at all places above p and $\overline{\rho}_{\pi,p} = \overline{\rho}$. Then there exists an automorphic rerpesentation π' such that $\rho = \rho_{\pi',p}$.*

We will now indicate some ways in which base change and the Jacquet-Langlands correspondence come up in the proof of this theorem. To begin with, we can use base change to make some immediate reductions that simplify the situation. For example, our representation ρ is of the form

$$\begin{pmatrix} \alpha\chi_p & * \\ & \beta \end{pmatrix}$$

at each place above p. By making a solvable base change, we can reduce to the case where α and β are unramified. Even more drastically, we can make a solvable base change to reduce to the case where $\overline{\rho}|_{G_{F_v}}$ is *trivial* at any given finite set of places. Moving to such a situation can make some of the local deformation theory easier. Two other things we can do with base change: we can reduce to the case where $\det \rho$ is the cyclotomic character (our hypotheses imply that it is a finite twist of the cyclotomic character); and we can reduce to the case that $F/\mathbf{Q}$ has even degree, which is useful for finding quaternion algebras with prescribed ramification.

The above applications of base change are very useful but fairly superficial. We now describe a more serious application. In the hypotheses of the theorem, we have been given an automorphic rerpesentation

π such that $\overline{\rho} = \overline{\rho}_{\pi,p}$. In the proof, however, we need it to be the case that ρ and $\rho_{\pi,p}$ are potentially unramified at the same set of places. This need not be the case for the π we have. Of course, we are free to replace π with another form π' such that $\overline{\rho}_{\pi,p} = \overline{\rho}_{\pi',p}$, that is, one that is congruent to π modulo p (while still maintaining the other hypotheses). So the question is: given π as in the theorem, can we find a congruent π' such that $\rho_{\pi',p}$ and ρ are potentially unramified at the same set of places? Alternatively, we know that $\rho_{\pi',p}$ is potentially unramified precisely at the places where it is not Steinberg, so we could also ask if we can replace π by a congruent form and prescribe the set of places at which this new form is Steinberg.

Clearly, this issue cannot be resolved with base change; in fact, it requires some real work. In the early days of the modularity lifting theorem, these congruences were found using the geometry of the modular curves. These proofs were difficult and fairly specific. Since then, new proofs have been found which are easier and more general. The common theme of these proofs is to use the Jacquet-Langlands correspondence and then do some computations with modular forms on quaternion algebras — which are just functions on a finite set. It is much easier to manipulate these functions than forms on the modular curve!

To prove the theorem we identify a cerain universal deformation ring of the Galois representation ρ with a certain Hecke algebra. Originally, this Hecke algebra was one for GL(2). However, by Jacquet-Langlands, we can find the same Hecke algebra on a quaternion algebra, and as we have explained, it is often easier to prove things in that setting. So we will in fact use a Hecke algebra on a quaternion algebra. Thus the Jacquet-Langlands correspondence will be built into our proof at a very fundamental level.

Lecture 18: Overview of the Taylor-Wiles method

Andrew Snowden
March 11, 2010

1. Statement of theorem

The goal of this lecture is to sketch a proof of the following modularity lifting theorem.

Theorem 1. *Let $F/\mathbf{Q}$ be a totally real number field and let $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ be a continuous representation of its absolute Galois group, with $p > 5$. Assume that ρ satisfies the following conditions:*

- *ρ ramifies at only finitely many places.*
- *ρ is odd, i.e., $\det \rho(c) = -1$ for all complex conjugations $c \in G_F$.*
- *ρ is potentially crystalline and ordinary at all places above p.*
- *$\overline{\rho}|_{G_{F(\zeta_p)}}$ is absolutely irreducible.*
- *There exists a parallel weight two Hilbert modular form f such that ρ_f is potentially crystalline and ordinary at all places above p and $\overline{\rho} = \overline{\rho}_f$.*

Then there exists a Hilbert modular form g such that $\rho = \rho_g$.

We will prove this theorem by proving an $R = \mathbf{T}$ theorem, where R is a deformation ring of $\overline{\rho}$ with certain local conditions imposed and $\mathbf{T}$ is a certain Hecke algebra. Clearly, if we have an appropriate $R = \mathbf{T}$ theorem then we get a modularity lifting theorem, as ρ defines a homomorphism $R \to \overline{\mathbf{Q}}_p$ and thus (by $R = \mathbf{T}$), a homomorphism $\mathbf{T} \to \overline{\mathbf{Q}}_p$, which is the same as a modular form.

2. Initial reductions

As we have previously explained, by using base change we may pass to totally real solvable extensions of F. The hypotheses of the theorem imply that $\det \rho$ is of the form $\chi_p \psi$ where ψ is a finite order character of G_F. It is not difficult to see that there is a totally real solvable extension F'/F such that $\psi|_{G_{F'}} = (\psi')^2$ for some finite order character ψ' of G_F. Thus $(\psi')^{-1}\rho_f|_{G_{F'}}$ has determinant χ_p. Since twisting by a character does not affect modularity, it is enough to show that $(\psi')^{-1}\rho|_{G_{F'}}$ is modular. We may therefore assume $\det \rho = \chi_p$. Similarly, after possibly another base change, we can assume that $\det \rho_f = \chi_p$ as well.

Let $v \nmid p$ be a place of F. Call a representation $G_{F_v} \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ *Steinberg* if it is of the form

$$\begin{pmatrix} \chi_p & * \\ & 1 \end{pmatrix}.$$

If $\rho_v : G_{F_v} \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ is any continuous representation, then there is a finite extension F_v'/F_v such that $\rho_v|_{G_{F_v'}}$ is either unramified or Steinberg. We may thus make a global solvable extension F'/F such that $\rho|_{G_{F'}}$ and $\rho_f|_{G_{F'}}$ are unramified or Steinberg at all places away from p. Let S be the set of places (away from p) at which ρ is ramified (and thus Steinberg), and let S' be the corresponding set for ρ_f. By making another solvable extension, we may assume that $\rho|_{G_{F_v}}$ and $\rho_f|_{G_{F_v}}$ are crystalline at $v \mid p$. Finally, we may pass to another solvable extension and assume that $\overline{\rho}|_{G_{F_v}}$ and $\overline{\rho}_f|_{G_{F_v}}$ are trivial at all places v above p or in $S \cup S'$.

Now, we will not be able to prove the strongest possible form of an $R = \mathbf{T}$ theorem. We must impose the following hypothesis: the local deformation spaces used to construct R must be connected. Practically speaking, this means that $\rho|_{G_{F_v}}$ and $\rho_f|_{G_{F_v}}$ must lie on the same irreducible component of the universal semi-stable deformation space of $\overline{\rho}|_{G_{F_v}}$. What are the components of this space? For $v \mid p$ there are three components: ordinary crystalline, ordinary non-crystalline and non-ordinary. Thus our "ordinary and crystalline" hypotheses ensure that there is no problem at the places above p. Unramified representations always lie on the same component, so there is no problem outside of $S \cup S'$. However, if $v \nmid p$ then the universal semi-stable deformation space of the trivial representation has two components: unramified and Steinberg. We must therefore assume $S = S'$, which is a non-trivial assumption that may not be satisfied by the ρ_f that is given to us. However, one can always find a congruence with a form f' which does satisfy this condition. (Finding this f' is not at all trivial, but occurs outside the scope of the $R = \mathbf{T}$ theorem, and we will not discuss it further in this lecture.)

By making a further base change, we may assume that F has even degree over $\mathbf{Q}$ and that S has even cardinality.

3. The $R = \mathbf{T}$ theorem: set-up

We begin by precisely stating the situation in which we have placed ourselves. We have a representation

$$\overline{\rho} : G_F \to \mathrm{GL}_2(k)$$

where k is a finite field of characteristic p, a finite set S of places of F away from p and a modular representation ρ_f lifting $\overline{\rho}$. Let S_p denote the places of F above p. We assume the following hypotheses:

(A1) ρ_f is crystalline and ordinary at all places in S_p, Steinberg at all places in S and unramified at all other places.
(A2) $\det \rho_f = \chi_p$.
(A3) $\overline{\rho}|_{G_{F(\zeta_p)}}$ is absolutely irreducible.
(A4) $\overline{\rho}|_{G_{F_v}}$ is trivial for $v \in S_p \cup S$.
(A5) F has even degree over $\mathbf{Q}$ and S has even cardinailty.

Note that the representation we had previously called ρ has completely disappeared from the set-up. It may not be absolutely necessary to assume (A4), but it does not cost us anything to do so, and makes study of the local deformation rings a bit easier.

We now define the ring R. Let $\widetilde{R}$ be the universal deformation ring of $\overline{\rho}$ unramified outside of $S \cup S_p$ and with determinant χ_p. (Here we take coefficients in some fixed $\mathscr{O}/\mathbf{Z}_p$ with residue field k.) For a place v let $\widetilde{R}_v$ be the universal deformation ring of $\overline{\rho}|_{G_{F_v}}$ with determinant χ_p. For $v \in S_p$ let R_v be the quotient of $\widetilde{R}_v$ classifying ordinary crystalline representations. For $v \in S$ let R_v be the quotient of $\widetilde{R}_v$ classifying Steinberg representations. We then let R be the tensor product of $\widetilde{R}$ with $\otimes R_v$ over $\otimes \widetilde{R}_v$. The latter tensor products are over $S \cup S_p$ and we should complete these tensor products.

We now define the Hecke algebra. Let D be the unique quaternion algebra over F ramifying at all the infinite places and the places in S (and nowhere else). This exists by (A5). For a compact open set U of $(D \otimes \mathbf{A}_F^f)^\times$ let $S_2(U)$ denote the set of functions

$$X(U) = D^\times \backslash (D \otimes \mathbf{A}_F^f)^\times / ((\mathbf{A}_F^f)^\times \cdot U) \to \mathscr{O}.$$

For places v at which U is maximal and D unramified, there is a Hecke operator T_v acting on $S_2(U)$.

Let U° be "the" maximal compact open subgroup of $(D \otimes \mathbf{A}_F^f)^\times$. Let $\mathbf{T}^{(p)}$ be the subalgebra of $\mathrm{End}(S_2(U^\circ))$ generated by the Hecke operators T_v for $v \notin S_p \cup S$. Let $\mathbf{T}$ be the subalgebra generated by the T_v for $v \notin S$. Thus $\mathbf{T}$ contains the Hecke operators above p and $\mathbf{T}^{(p)}$ does not. By the Jacquet-Langlands correspondence and conditions (A1) and (A2), our modular form f can be transferred to an element of $S_2(U^\circ)$ which is an eigenform for $\mathbf{T}$. The form f defines a homomorphism $\mathbf{T} \to \mathscr{O}$, the kernel of which is contained in a unique maximal ideal $\mathfrak{m}$. The form f is actually irrelevant; all that matters is the ideal $\mathfrak{m}$. It has the following two properties (which characterize it uniquely):

- For $v \notin S \cup S_p$, the image of T_v in $\mathbf{T}/\mathfrak{m}$ is equal to $\mathrm{tr}\,\overline{\rho}(\mathrm{Frob}_v)$.
- For $v \in S_p$, the Hecke operator T_v does not belong to $\mathfrak{m}$.

The first condition means that $\mathfrak{m}$ is associated to the representation $\overline{\rho}$; the second is the ordinarity condition. We regard $\mathfrak{m}$ as an ideal of $\mathbf{T}^{(p)}$ by contraction. (Remark: we need the group U° to satisfy a certain smallness condition which our group U° does not. To get a correct definition of U° one picks an auxiliary place v_{aux} with certain nice properties and takes $U^\circ_{v_{\mathrm{aux}}}$ to be sufficiently small; away from v_{aux} the group U° is maximal. One must also modify the definition of $\widetilde{R}$ to allow for ramification at v_{aux}. We will ignore this subtlety for now.)

As we have seen, there is a representation

$$G_F \to \mathrm{GL}_2(\mathbf{T}^{(p)}_{\mathfrak{m}}),$$

which lifts $\overline{\rho}$. This induces a surjection $\widetilde{R} \to \mathbf{T}^{(p)}_{\mathfrak{m}}$. The above representation is Steinberg at all the places in S (since D ramifies at S and local and global Langlands are compatible). However, it is not necessarily ordinary at the places in S_p (though it is automatically crystalline). This should not be surprising, as we

have not told $\mathbf{T}_{\mathfrak{m}}^{(p)}$ anything about what is happening at p. There is a map $\mathbf{T}_{\mathfrak{m}}^{(p)} \to \mathbf{T}_{\mathfrak{m}}$; composing this with the above representation gives a representation

$$G_F \to \mathrm{GL}_2(\mathbf{T}_{\mathfrak{m}}).$$

This representation *is* ordinary at the places above p; one should think that the Hecke operators at p specify the local component of the deformation space at p. Thus the map $\widetilde{R} \to \mathbf{T}_{\mathfrak{m}}$ factors through R. Unfortunately, this map is no longer surjective. However, it is not very far from being surjective and the problem can be controlled locally at p: for $v \in S_p$ there is a modified local deformation ring R'_v which is a finite R_v-algebra and isomorphic to R_v after p is inverted. Define R' to be like R but use R'_v for $v \in S_p$. Then there is a natural surjection $R' \to \mathbf{T}_{\mathfrak{m}}$. Our goal is to prove:

Theorem 2. *The map $R'[1/p] \to \mathbf{T}_{\mathfrak{m}}[1/p]$ is an isomorphism and R' is finite over $\mathscr{O}$.*

The ring $\mathbf{T}_{\mathfrak{m}}$ is torsion-free by consturction. This theorem does not allow us to control the torsion in R', except to say that it is finite. One expects that R' is torsion free; this may actually be proved in the case we are in (the ordinary case), but I do not know for certain. One can modify the proof of the above theorem to show that R is finite over $\mathscr{O}$, which is often more relevant but does not seem to follow formally from the finiteness of R'.

4. Taylor-Wiles primes

The basic idea to the proof of Theorem 2 (called the Taylor-Wiles method) is to find a tower of maps $R_n \to T_n$ lifting $R' \to \mathbf{T}_{\mathfrak{m}}$ and then build a sort of inverse limit R_∞ and T_∞ out of the R_n and T_n, which is a nice ring. We can then prove that $R_\infty[1/p] \to T_\infty[1/p]$ is an isomorphism and deduce from this the statement we want. Actually, the Hecke algebras will not be so important; they will be replaced by spaces of modular forms.

We find the rings R_n be introducing certain auxiliary deformation rings. By a *TW set of places* we mean a finite set Q of places of F satisfying the following conditions:

- Q is disjoint from S_p and S. (And does not contain v_{aux}.)
- $\mathbf{N}\, v = 1 \pmod{p}$ for all $v \in Q$.
- The eigenvalues of $\overline{\rho}(\mathrm{Frob}_v)$ are distinct and belong to k.

(The "belong to k" part is not serious — we can replace k by its quadratic extension and then all eigenvalues of all Frobenii belong to k.) Given such a set Q of places, we define R_Q similarly to R' except we allow ramification at the places in Q. Note that $R_\emptyset = R'$. We will typically write $R_\emptyset$ in place of R' for notational uniformity.

Although we did not impose any deformation conditions at the places in Q, the conditions on the places in Q has strong consequences. Let $v \in Q$. Then the universal deformation $G_F \to \mathrm{GL}_2(R_Q)$ restricted to G_{F_v} is a sum of two characters $\eta_1 \oplus \eta_2$. These characters are necessarily tamely ramified, since their reduction is unramified, and in fact the image of inertia is p-power. By class field theory, η_1 defines a map $F_v^\times \to R_Q^\times$. We thus get a map $(\mathscr{O}_{F_v}/\mathfrak{p}_v)^\times \to R_Q^\times$, which factors through the maximal p-power quotient of $(\mathscr{O}_{F_v}/\mathfrak{p}_v)^\times$. Define Δ_Q to be the product of the maximal p-power quotients of $(\mathscr{O}_{F_v}/\mathfrak{p}_v)^\times$ for $v \in Q$. Then we have just given R_Q the structure of an $\mathscr{O}[\Delta_Q]$-algebra. This was not quite canonical, since we had to choose η_1. Define a *TW datum* to be a pair $(Q, \{\alpha_v\})$ where Q is a TW set of primes and α_v is a chosen eigenvalue of $\overline{\rho}(\mathrm{Frob}_v)$ for each place $v \in Q$. Given such a datum we get a canonical $\mathscr{O}[\Delta_Q]$-algebra structure on R_Q, since we can take η_1 to correspond to α_v. The following result is not difficult:

Proposition 3. *The canonical map $R_Q \to R_\emptyset$ is surjective. Its kernel is $\mathfrak{a}_Q R_Q$, where $\mathfrak{a}_Q$ is the augmentation ideal of $\mathscr{O}[\Delta_Q]$.*

We now define the auxiliary Hecke algebras and spaces of modular forms. Thus let $(Q, \{\alpha_v\})$ be a TW-datum. We define compact open subgroups $U_Q \subset V_Q$. At places $v \notin Q$, we define $U_{Q,v} = V_{Q,v} = U_v^\circ$. Let $v \in Q$. We then define

$$V_{Q,v} = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathscr{O}_{F_v}) \;\middle|\; c \in \mathfrak{p}_v \right\}$$

and

$$U_{Q,v} = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathscr{O}_{F_v}) \;\middle|\; c \in \mathfrak{p}_v \text{ and } ad^{-1} \text{ maps to 1 in } \Delta_Q \right\}.$$

Of course, $V_\emptyset = U_\emptyset = U^\circ$. Note that U_Q is normal in V_Q and V_Q/U_Q is identified with Δ_Q.

Let $\mathbf{T}(V_Q)$ be the subalgebra of $\mathrm{End}(S_2(V_Q))$ generated by the T_v for $v \notin S \cup Q$, and let $\mathbf{T}_+(V_Q)$ be the subalgebra generated by these T_v together with the U_v for $v \in Q$. We have a map $\mathbf{T}(V_Q) \to \mathbf{T}(V_\emptyset)$ and can thus regard $\mathfrak{m}$ as an ideal of $\mathbf{T}(V_Q)$. We also have a map $\mathbf{T}(V_Q) \to \mathbf{T}_+(V_Q)$. Let $\mathfrak{n}_Q$ be the ideal of $\mathbf{T}_+(V_Q)$ generated by $\mathfrak{m}$ and $U_v - \alpha_v$ for $v \in Q$. We then have the following result:

Proposition 4. *We have an isomorphism $S_2(V_Q)_{\mathfrak{n}_Q} = S_2(U^\circ)_\mathfrak{m}$.*

Proof. We just indicate the map $S_2(U^\circ)_\mathfrak{m} \to S_2(V_Q)_{\mathfrak{n}_Q}$ in the case that $Q = \{v\}$ (it suffices to treat this case by an inductive argument). By Hensel's lemma we have a factorization

$$X^2 - T_v X + \mathbf{N}\, v = (X - A)(X - B)$$

for A and B in $\mathbf{T}(U^\circ)_\mathfrak{m}$, with A mapping to α_v modulo $\mathfrak{m}$. We thus have a map

$$S_2(U^\circ)_\mathfrak{m} \to S_2(V_Q), \qquad f \mapsto Af - \begin{pmatrix} 1 & \\ & \varpi_v \end{pmatrix} f.$$

Here ϖ_v denotes a uniformizer at v and the operator U_v is defined using the double coset of $\begin{pmatrix} \varpi_v & \\ & 1 \end{pmatrix}$. One must show that the above map actually lands in $S_2(V_Q)_{\mathfrak{n}_Q}$ (which we think of as a summand of $S_2(V_Q)$) and that it is an isomorphism. Note in particular, that the proposition implies that $S_2(V_Q)_{\mathfrak{n}_Q}$ consists only of old forms. We note that there is also an isomorphism $\mathbf{T}_+(V_Q)_{\mathfrak{n}_Q} = \mathbf{T}_\mathfrak{m}$, though we will not need it. □

Let $\mathbf{T}_+(U_Q)$ be the subalgebra of $\mathrm{End}(S_2(U_Q))$ generated by the T_v for $v \notin S \cup Q$ and the U_v for $v \in Q$. There is a natural map $\mathbf{T}_+(U_Q) \to \mathbf{T}'_+(V_Q)$. We let $\mathfrak{m}_Q$ be the contraction of $\mathfrak{n}_Q$ under this map. We let T_Q be the localization $\mathbf{T}_+(U_Q)_{\mathfrak{m}_Q}$ and we let M_Q be the localization $S_2(U_Q)_{\mathfrak{m}_Q}$. Note that $T_\emptyset = \mathbf{T}_\mathfrak{m}$ and that $V_Q/U_Q = \Delta_Q$ acts on M_Q. We then have

Proposition 5. *The space M_Q is free over $\mathscr{O}[\Delta_Q]$. There is a natural isomorphism $M_Q/\mathfrak{a}_Q M_Q \to M_\emptyset$.*

Proof. The map $X(U_Q) \to X(V_Q)$ is a Galois cover with group Δ_Q. (This uses the smallness hypothesis on U°.) From this, one easily deduces that M_Q is free over $\mathscr{O}[\Delta_Q]$ and that $M_Q/\mathfrak{a}_Q M_Q$ is naturally isomorphic to $S_2(V_Q)_{\mathfrak{n}_Q}$. To get the isomorphism with $M_\emptyset$ apply the previous proposition. □

We have a Galois representation $G_F \to \mathrm{GL}_2(T_Q^{(p)})$, which yields a surjection $\widetilde{R}_Q \to T_Q^{(p)}$ (where $\widetilde{R}_Q$ is the universal deformation ring of $\overline{\rho}$ unramified outside $S \cup S_p \cup Q$). This Galois representation is not necessarily ordinary at the places above p, but the induced representation $G_F \to \mathrm{GL}_2(T_Q)$ is. The resulting map $\widetilde{R}_Q \to T_Q$ is not longer surjective, but there is a natural surjection $R_Q \to T_Q$ of $\widetilde{R}_Q$-algebras.

In the next section, it will be important to have framed versions of everything. For $v \in S \cup S_p$ we let $R_v^\square$ be the framed local deformation ring. (It is actually the only one that makes sense; what we had called R_v before does not really exist as a ring.) We let $(R'_v)^\square$ be the modification at places above p. We let B be the tensor product of the $R_v^\square$ for $v \in S$ and the $(R'_v)^\square$ for $v \in S_p$. We let $R_Q^\square$ be like R_Q but have framings at all places in $S \cup S_p$. It is an algebra over B. Finally, we define $T_Q^\square = T_Q \otimes_{R_Q} R_Q^\square$ and $M_Q^\square = M_Q \otimes_{T_Q} T_Q^\square$.

5. The patching argument

For each TW set of primes we have constructed a ring R_Q and given it the structure of a module over $\mathscr{O}[\Delta_Q]$, which is a ring of the form $\mathscr{O}[T_1, \dots, T_n]/(T_i^{p^{a_i}} = 1)$ where $n = \#Q$. We would like to use various Q's to build a ring R_∞ which is an algebra over $\mathscr{O}[\![T_1, \dots, T_n]\!]$ for some n (which does not obviously factor through a large quotient). To do this, we need to hold $\#Q$ fixed and let its elements have norm congruent to 1 modulo higher and higher powers of p. There will not be natural maps between the various R_Q's, but we will nonetheless manage to find maps between pieces of these rings by a sort of pigeonhole principle.

We now assume that we can find integers h and g and for each n a TW set of primes Q_n satisfying the following conditions:

- $\#Q_n = h$
- $\mathbf{N}\, v = 1 \pmod{p^n}$ for all $v \in Q_n$
- $R_{Q_n}^\square$ is topologically generated by g elements over B.

We extend each Q_n to a TW datum by arbitrarily choosing eigenvalues. We write $R_n^{\square}$ in place of $R_{Q_n}^{\square}$, and make this convention for other notations (e.g., $M_n^{\square}$). We did not give motivation for the last condition, but it is a natural condition to impose if we want patched ring $R_\infty^{\square}$ to be finitely generated. We fix surjections $B[\![z_1,\dots,z_g]\!] \to R_n^{\square}$ for each n.

For a complete local ring A let $\mathfrak{m}_A$ denote its maximal ideal. Also, let $\mathfrak{m}_A^{(n)}$ denote the ideal generated by nth powers of elements of A; this is not the same as $\mathfrak{m}_A^n$. Let $\ell = 4(\#S_p + \#S) - 1$ and define

$$P = \mathscr{O}[\![x_1,\dots,x_\ell,y_1,\dots,y_h]\!].$$

We make $R_n^{\square}$ into a P-algebra by letting the x_i be the framing variables and letting the y_i act through a chosen surjection $\mathscr{O}[\![y_1,\dots,y_h]\!] \to \mathscr{O}[\Delta_n]$. For an integer n let $\mathfrak{c}_n$ be the ideal of P generated by

$$(\pi^n, x_1^{p^n},\dots,x_\ell^{p^n},(y_1+1)^{p^n}-1,\dots,(y_h+1)^{p^n}-1)$$

(where π is a uniformizer of $\mathscr{O}$). Let s denote the rank of M_0 over $\mathscr{O}$. For an integer n let $r_n = snp^n(h+\ell)$. We remind the reader that $R_n^{\square}$ is an algebra over B and that $M_n^{\square}$ is an $R_n^{\square}$-module.

A *patching datum of level n* consists of the following:

- A complete local B-algebra D with $\mathfrak{m}_D^{(r_n)} = 0$.
- A map of $\mathscr{O}$-algebras $P/\mathfrak{c}_n \to D$.
- A surjection of B-algebras $D \to R_0^{\square}/(\mathfrak{c}_n R_0^{\square} + \mathfrak{m}_{R_0^{\square}}^{(r_n)})$.
- A surjection of B-algebras $B[\![z_1,\dots,z_g]\!] \to D$.
- A D-module L which is finite free over $P/\mathfrak{c}_n$ of rank s.
- A surjection of $B[\![z_1,\dots,z_g]\!]$ modules $L \to M_0^{\square}/\mathfrak{c}_n M_0^{\square}$.

The number of elements of D is finite (it can be bounded in terms of B and n). We thus find that, up to the obvious notion of isomorphism, there are only finitely many patching data of a given level.

Let $m \geq n$ be integers. Put

$$D_{n,m} = R_m^{\square}/(\mathfrak{c}_n R_m^{\square} + \mathfrak{m}_{R_m^{\square}}^{(r_n)}), \qquad L_{n,m} = M_m^{\square}/\mathfrak{c}_n M_m^{\square}.$$

One verifies that $(D_{n,m}, L_{n,m})$ is a patching datum of level n. Since there only only finitely many patching data of a given level, we can pass to a subsequence and assume $D_{n,m} = D_{n,n}$ and $L_{n,m} = L_{n,n}$ for all $m \geq n$. Denote the common value by D_n and L_n. Then the maps

$$D_{n+1}/(\mathfrak{c}_n D_{n+1} + \mathfrak{m}_{D_{n+1}}^{(r_n)}) \to D_n, \qquad L_{n+1}/\mathfrak{c}_n L_n \to L_n$$

are isomorphisms.

Let $R_\infty^{\square}$ be the inverse limit of the D_n and $M_\infty^{\square}$ the inverse limit of the L_n. The space $M_\infty^{\square}$ is a free P-module of rank s. The ring $R_\infty^{\square}$ is a P-algebra and a B-algebra, and there is a given surjection

$$B[\![z_1,\dots,z_g]\!] \to R_\infty^{\square}.$$

Since P is a power series ring, the map $P \to R_\infty^{\square}$ can be lifted through the above surjection. We now have the following lemma:

Lemma 6. *Let $R \to S$ be a map of noetherian domains of the same dimension and let M be a non-zero S-module which is finite projective over R. Then $R \to S$ is a finite map. If R and S are regular then M is a finite projective faithful S-module.*

Now, by the way we chose our deformation conditions, B is a domain and $B[1/p]$ is smooth over $\mathbf{Q}_p$. (These are theorems that we need to prove!) Note that B being a domain is the hypothesis mentioned at the beginning of these notes, that our local deformation spaces needed to be irreducible. We now assume:

$$\dim B = 1 + h + \ell - g.$$

We will address this assumption below. This dimension assumption implies that P and $B[\![z_1,\dots,z_g]\!]$ have the same dimension. We conclude from the lemma that $B[\![z_1,\dots,z_g]\!]$ is finite over P and $M[1/p]$ is a faithful $B[\![z_1,\dots,z_g]\!][1/p]$ module. The former implies that $R_\infty^{\square}$ is finite over P while the latter implies that $M_\infty^{\square}[1/p]$ is a faithful $R_\infty^{\square}[1/p]$ module (since the map $B[\![z_1,\dots,z_g]\!] \to \mathrm{End}(M_\infty^{\square})$ factors through $R_\infty^{\square}$).

Now, by the construction of $R_\infty^{\square}$ and $M_\infty^{\square}$ we have isomorphisms

$$R_\infty^{\square}/(y_1,\dots,y_h)R_\infty^{\square} \to R_0^{\square}, \qquad M_\infty^{\square}/(y_1,\dots,y_h)M_\infty^{\square} \to M_0^{\square}.$$

It follows that $R_0^\square$ is finite over $\mathscr{O}[\![x_1, \dots, x_\ell]\!]$, which implies that R_0 is finite over $\mathscr{O}$. Since $M_0^\square$ is free over P, we see that the action of $R_0^\square[1/p]$ on $M_0^\square[1/p]$ is still faithful. Since this action came via the map $R_0^\square \to T_0^\square$, we conclude that $R_0^\square[1/p] \to T_0^\square[1/p]$ is injective. Since we already knew this to be surjective, it must be an isomorphism.

6. Resolving the assumptions

In the last section we proved Theorem 2 assuming the following: there exist integers h and g satisfying

$$\dim B = 1 + h + \ell - g$$

such that for every integer n there is a set of primes Q_n satisfying:

- $\#Q_n = h$
- $\mathbf{N}\, v = 1 \pmod{p^n}$ for all $v \in Q_n$
- R_{Q_n} is topologically generated by g elements over B.

In fact, one can find Q_n as above with

$$h = \dim H^1(G_{F,S}, \mathrm{ad}^\circ \overline{\rho}(1))$$
$$g = h - [F : \mathbf{Q}] + \#S + \#S_p - 1.$$

The proof of this will probably require its own talk; it is purely Galois theoretic and makes no use of modular forms. It uses condition (A3), the assumption $p > 5$ and certain conditions on v_{aux} that we did not state. Now,

$$\dim R_v - \dim \mathscr{O} = \begin{cases} 3 & v \in S \\ 3 + [F_v : \mathbf{Q}_p] & v \in S_p \end{cases}$$

and so

$$\begin{aligned} \dim B - \dim \mathscr{O} &= \sum_{v \in S} 3 + \sum_{v \in S_p} (3 + [F_v : \mathbf{Q}_p]) \\ &= 3\#S + 3\#S_p + [F : \mathbf{Q}] \\ &= h + \ell - g. \end{aligned}$$

(Since B is the tensor product of the R_v over $\mathscr{O}$, the *relative* dimension of B over $\mathscr{O}$ is the sum of the relative dimensions of the R_v over $\mathscr{O}$.) Therefore everything works!

REVIEW OF GALOIS DEFORMATIONS

SIMON RUBINSTEIN-SALZEDO

1. Deformations and framed deformations

We'll review the study of Galois deformations. Here's the setup. Let G be a profinite group, and let $\overline{\rho} : G \to \mathrm{GL}_n(k)$ be a representation of G over a finite field k of characteristic p. Let Λ be a complete DVR with residue field k, and let $\mathcal{C}_\Lambda$ denote the category whose objects are artinian local Λ-algebras with residue field k, and whose morphisms are local homomorphisms. Let $\widehat{\mathcal{C}}_\Lambda$ denote the category of complete noetherian local Λ-algebras with residue field k, which is the completion of $\mathcal{C}_\Lambda$ under limits. Frequently, Λ will be $W(k)$, the ring of Witt vectors over k.

We now define two deformation functors associated to $\overline{\rho}$. The first is the deformation functor $\mathrm{Def}(\overline{\rho}) : \widehat{\mathcal{C}}_\Lambda \to \underline{\mathrm{Sets}}$ given by

$$\mathrm{Def}(\overline{\rho})(A) = \{(\rho, M, \phi)\}/\cong,$$

where M is a free A-module of rank n, $\rho : G \to \mathrm{Aut}_A(M)$, and $\phi : \rho \otimes_A k \to \overline{\rho}$ is an isomorphism. The second is the framed deformation functor $\mathrm{Def}^\square(\overline{\rho}) : \widehat{\mathcal{C}}_\Lambda \to \underline{\mathrm{Sets}}$ given by

$$\mathrm{Def}^\square(\overline{\rho})(A) = \{(\rho, M, \phi, \mathcal{B}) \mid (\rho, M, \phi) \in \mathrm{Def}(\overline{\rho})(A)\}/\cong,$$

where $\mathcal{B}$ is a basis of M which is sent to the standard basis for k^n under ϕ.

We can compute both Def and $\mathrm{Def}^\square$ at the level of its artinian quotients: if $\mathfrak{m}$ is the maximal ideal of A, then

$$\mathrm{Def}(\overline{\rho})(A) = \varprojlim \mathrm{Def}(\overline{\rho})(A/\mathfrak{m}^i),$$
$$\mathrm{Def}^\square(\overline{\rho})(A) = \varprojlim \mathrm{Def}^\square(\overline{\rho})(A/\mathfrak{m}^i).$$

The functors Def and $\mathrm{Def}^\square$ are not always representable. However, we impose some restrictions to guarantee that at least $\mathrm{Def}^\square$ will be. We say that G satisfies the p-finiteness condition if for every open subgroup $H \subset G$ of finite index, there are only finitely many continuous homomorphisms $H \to \mathbb{Z}/p\mathbb{Z}$. From now on, we'll assume that G satisfies the p-finiteness condition. This is a reasonable assumption, since it holds in cases we're likely to care about. For example, if K is a global field not of characteristic p and S is a finite set of places, then $G_{K,S}$ satisfies the p-finiteness condition for all p. Also, if K is a local field of residue characteristic p, then G_K satisfies the ℓ-finiteness condition for all $\ell \neq p$.

Date: 30 March, 2010.

In this case, $\mathrm{Def}^{\square}(\overline{\rho})$ is representable; call its representing object $R^{\square}(\overline{\rho})$. If $\mathrm{Def}(\overline{\rho})$ is also representable, call its representing object $R(\overline{\rho})$. These are the framed deformation ring and the deformation ring, respectively. Concretely, when $\mathrm{Def}^{\square}(\overline{\rho})$ or $\mathrm{Def}(\overline{\rho})$ is representable, this means that there is some ring $R^{\square}(\overline{\rho})$ or $R(\overline{\rho})$ so that any deformation factors uniquely through the map

$$G \to \mathrm{GL}_n(R^{\square}(\overline{\rho})) \text{ or } G \to \mathrm{GL}_n(R(\overline{\rho})).$$

Schlessinger's criterion tells us when functors $\mathcal{C}_\Lambda \to \underline{\mathrm{Sets}}$ are (pro)-representable. When we apply it to the case of the deformation functor, we get the following:

Proposition 1. *If G satisfies the p-finiteness condition and $\mathrm{End}_G(\overline{\rho}) = k$ ($\overline{\rho}$ is absolutely irreducible, meaning that $\overline{\rho} \otimes_k k'$ is irreducible for all finite extensions k'/k), then $\mathrm{Def}(\overline{\rho})$ is representable.*

For another example of representability of deformation functors, we review ordinary defomrations. Let K be a p-adic field, and let $\psi : G_K \to \mathbb{Z}_p^\times$ be the p-adic cyclotomic character. An n-dimensional representation ρ of G is said to be (distinguished) ordinary if there exist integers $e_1 > e_2 > \cdots > e_{n-1} > e_n = 0$ so that

$$\rho\,|_{I_K} \sim \begin{pmatrix} \psi^{e_1} & * & \cdots & * \\ 0 & \psi^{e_2} & \cdot & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & \cdots & \psi^{e_n} = 1 \end{pmatrix}.$$

For fixed $e_1, \ldots, e_n$ which are distinct modulo $p-1$, the ordinary deformation functor $\mathrm{Def}^{\mathrm{ord}}(\overline{\rho})$ is the subfunctor of $\mathrm{Def}(\overline{\rho})$ consisting of only the distinguished ordinary lifts of $\overline{\rho}$ for that choice of $e_1, \ldots, e_n$.

Proposition 2. *If a two-dimensional residual representation $\overline{\rho}$ is non-split, then $\mathrm{Def}^{\mathrm{ord}}(\overline{\rho})$ is representable. More generally, if $\overline{\rho}$ is n-dimensional, and every 2×2 diagonal minor is non-split, then $\mathrm{Def}^{\mathrm{ord}}(\overline{\rho})$ is representable.*

Let's recall some properties of deformation rings. An important property of deformation rings is that they commute with finite extension of residue fields. That is, suppose k'/k is a finite extension. Let $\overline{\rho}$ be an absolutely irreducible residual representation of k; we can extend scalars to k' to get an absolutely irreducible residual representation of k'. Then, if $W(k)$ denotes the ring of Witt vectors over k, we have

$$R^{\square}(\overline{\rho}) \otimes_{W(k)} W(k') \cong R^{\square}(\overline{\rho} \otimes_k k'),$$
$$R(\overline{\rho}) \otimes_{W(k)} W(k') \cong R(\overline{\rho} \otimes_k k').$$

Suppose $\overline{\rho}$ is an absolutely irreducible residual representation of dimension N, and let $\det : \mathrm{GL}_N(k) \to \mathrm{GL}_1(k)$ be the determinant map. Then there is a natural homomorphism

$$R(\det(\overline{\rho})) \to R(\overline{\rho}).$$

More generally, if $\delta : \mathrm{GL}_N \to \mathrm{GL}_M$ is a homomorphism of group schemes, we get a natural map of deformation rings

$$R(\delta(\overline{\rho})) \to R(\overline{\rho}).$$

Deformation rings also commute with tensor products of representations. Let $\overline{\pi}, \overline{\rho}$ be two absolutely irreducible residual representations whose tensor product is also absolutely irreducible. Then we get a natural map

$$R(\overline{\pi} \otimes \overline{\rho}) \to R(\overline{\pi}) \widehat{\otimes} R(\overline{\rho}).$$

If $\overline{\pi}$ is a one-dimensional representation, i.e. a character, we call this map **twisting** by $\overline{\pi}$.

Deformation rings are also functorial in the choice of profinite group. Let $\phi : G \to G'$ be a group homomorphism, and let $\overline{\rho}$ be a residual representation of G'. Then composition with ϕ gives a residual representation of G. This gives us a map

$$R_G(\overline{\rho}) \to R_{G'}(\overline{\rho}).$$

An important example of deformations comes from looking at the Zariski tangent space. Let $k[\varepsilon]$ denote the dual numbers. If $F : \widehat{\mathcal{C}}_\Lambda \to \underline{\mathrm{Sets}}$ is a functor, its tangent space is $F(k[\varepsilon]) =: t_F$.

Let $V \in t_{\mathrm{Def}(\overline{\rho})}$. Then $V/\varepsilon V \cong \overline{\rho}$, so we have a short exact sequence

$$0 \to \varepsilon V \to V \to \overline{\rho} \to 0.$$

As G-modules, $\varepsilon V \cong \overline{\rho}$, so

$$t_{\mathrm{Def}(\overline{\rho})} \cong \mathrm{Ext}^1_{k[G]}(\overline{\rho}, \overline{\rho}) = H^1(G, \mathrm{Ad}(\overline{\rho})) = (\mathfrak{m}/(\mathfrak{m}, p))^2$$

for $\overline{\rho}$ absolutely irreducible. Here, $\mathrm{Ad}(\overline{\rho})$ is defined as follows: it is the representation of G whose underlying vector space is $M_N(k)$, and whose G-action is given by $g.m = \overline{\rho}(g)^{-1} m \overline{\rho}(g)$.

Let G be G_K if K is a local field, or $G_{K,S}$ for some finite set of places S if K is a global field. Fix a residual representation $\overline{\rho} : G \to \mathrm{GL}_n(k)$. If a deformation functor F for $\overline{\rho}$ is represented by R, then we have

$$t_F = F(k[\varepsilon]) = \mathrm{Hom}_\Lambda(R, k[\varepsilon]) = \mathrm{Hom}_\Lambda(R/(\mathfrak{m}_R^2 + \mathfrak{m}_\Lambda), k[\varepsilon]).$$

Since we also have

$$\frac{R}{\mathfrak{m}_R^2 + \mathfrak{m}_\Lambda} = k \oplus \frac{\mathfrak{m}_R}{\mathfrak{m}_R^2 + \mathfrak{m}_\Lambda},$$

and the second summand has square zero, we have

$$t_F = \mathrm{Hom}_k\left(\frac{\mathfrak{m}_R}{\mathfrak{m}_R^2 + \mathfrak{m}_\Lambda}, k\right) = t_R^*,$$

where for $A \in \widehat{\mathcal{C}}_\Lambda$, we set

$$t_A = \frac{\mathfrak{m}_A}{\mathfrak{m}_A^2 + \mathfrak{m}_\Lambda},$$

and t_A^* is its dual.

2. Why study Galois deformations?

We're mostly interested in the case of $G = G_{K,S}$ for some finite set of places S of a number field K. There are several reasons that this is a good idea.

(1) We can specify a residual representation $\overline{\rho} : G_{K,S} \to \mathrm{GL}_N(k)$ using only a finite amount of data, and there are only finitely many such representations, for K, S, N, and k fixed. When $\overline{\rho}$ is absolutely irreducible, we saw that we have a universal deformation, so all the lifts of $\overline{\rho}$ can be packaged together into a single complete noetherian local ring with residue field k.
(2) We might sometimes be interested in studying those deformations of $\overline{\rho}$ that have particular properties. For example, we mentioned ordinary deformations earlier. Another possibility that's relevant to us would be to look at modular deformations: those representations coming from modular forms. These correspond to quotients of the universal deformation ring. Placing such conditions on the representations at least conjecturally amounts to imposing local conditions at the ramified primes. We'll discuss this a bit more later.

3. Galois cohomology

We now review some cohomology of local fields. Let K be a finite extension of $\mathbb{Q}_p$, with Galois group G_K. Let μ be the roots of unity of K^s. If M is a finite G_K-module, set $M' = \mathrm{Hom}(M, \mu)$. Then for $0 \leq i \leq 2$, the cup product

$$H^i(K, M) \otimes H^{2-i}(K, M') \to H^2(K, \mu) \cong \mathbb{Q}/\mathbb{Z}$$

is a perfect pairing.

We have a similar statement when M is an ℓ-adic representation of G_K. Let ℓ be a prime, possibly equal to p, and let F be a finite extension of $\mathbb{Q}_\ell$. Suppose T is a free $\mathfrak{o}_F$-module with a continuous $\mathfrak{o}_F$-linear G_K-action, and let $V = T \otimes_{\mathfrak{o}_F} F$ be the corresponding $\mathbb{Q}_\ell$-vector space. Then G_K acts on V as well. Let V^* be the dual representation given by $V^* = \mathrm{Hom}_{\mathfrak{o}_F}(V, F(1))$. Then we have the following duality induced by the cup product: for $0 \leq i \leq 2$,

$$H^i(K, V) \otimes H^{2-i}(K, V^*) \to H^2(K, F(1)) \cong F$$

is a perfect pairing.

If M is a finite G_K-module, then $H^i(K, M)$ is a finite group for $0 \leq i \leq 2$, and $H^i(K, M) = 0$ for $i \geq 3$. Let $h^i(K, M)$ be the size of $H^i(K, M)$. Then, we define the Euler-Poincaré characteristic of M to be

$$\chi(M) = \frac{h^0(K, M) h^2(K, M)}{h^1(K, M)}.$$

We recall a few key properties of the Euler-Poincaré characteristic.

- If $0 \to M'' \to M \to M' \to 0$ is a short exact sequence of finite G_K-modules, then $\chi(M'')\chi(M') = \chi(M)$.
- If $(p, \#M) = 1$, then $\chi(M) = 1$.
- More generally, if $x \in \mathfrak{o}_K$, let $\|x\|_K$ be the normalized absolute value of x, so that
$$\|x\|_K = \frac{1}{(\mathfrak{o}_K : x\mathfrak{o}_K)}.$$
If $\#M = n$, then
$$\chi(M) = \|n\|_K = p^{-[K:\mathbb{Q}_p]\,\mathrm{ord}_p(n)}.$$

Sometimes, we wish to talk about Euler-Poincaré characteristics when the G_K-module is a $\mathbb{Q}_p$-vector space or free $\mathbb{Z}_p$-module. In that case, it would not make sense to talk about the sizes of the cohomology groups, but only about their ranks or dimensions. We can make sense of this in the case of finite modules instead, by talking of their ranks as $\mathbb{F}_p$-vector spaces. Let's write $\tilde{h}^i(K, M)$ for the dimension of $H^i(K, M)$ over $\mathbb{F}_p$, and write
$$\tilde{\chi}(M) = \tilde{h}^0(K, M) - \tilde{h}^1(K, M) + \tilde{h}^2(K, M).$$
Then the above result is equivalent to
$$\tilde{\chi}(M) = -[K : \mathbb{Q}_p]\,\mathrm{ord}_p(n).$$
If M is instead a $\mathbb{Q}_p$-vector space or a free $\mathbb{Z}_p$-module, we'll let $\tilde{h}^i(K, M)$ be the dimension or rank of $H^i(K, M)$ as a $\mathbb{Q}_p$-vector space or $\mathbb{Z}_p$-module, and we'll define the Euler-Poincaré characteristic similarly.

Let V now be a $\mathbb{Q}_p$-vector space of dimension d. Find inside V a G_K-stable lattice T. Since $T = \varprojlim T/p^rT$ and cohomology commutes with inverse limits, we have
$$H^i(K, T) = \varprojlim H^i(K, T/p^rT) \cong \varprojlim H^i(K, M_r),$$
where M_r is a G_K-module of size p^{dr}. By the above,
$$\tilde{\chi}(M_r) = -dr[K : \mathbb{Q}_p].$$
Taking inverse limits gives us
$$\tilde{\chi}(T) = -d[K : \mathbb{Q}_p].$$
Tensoring doesn't change the dimensions of the cohomology groups, so we also have
$$\tilde{\chi}(V) = -d[K : \mathbb{Q}_p].$$

4. Examples of Deformation Rings

It can be helpful to have a rough idea of what deformation rings look like. When they exist, they tend to be quotients of power series rings over $\mathbb{Z}_p$. Let's look at some examples.

Let S be a finite set of places of $\mathbb{Q}$, and let $\overline{\rho} : G_{\mathbb{Q},S} \to \mathrm{GL}_2(\mathbb{F}_p)$ be a representation. Let E be the fixed field of $\ker(\overline{\rho})$, and let $H = \mathrm{Gal}(E/\mathbb{Q})$. Let

$$V = \mathrm{coker}\left(\mu_p(E) \to \bigoplus_{v \in S} \mu_p(E_v)\right).$$

A $k[H]$-module W is said to be **prime-to-adjoint** if there is some subgroup A of H of order prime to p so that W and the adjoint $k[H]$-module M are relatively prime as $k[A]$-modules (so they share no common ireducible subrepresentations as A-modules).

Let Z_S be the set of $x \in \mathbb{Q}^\times$ so that (x) is a p^{th} power, and so that $x \in E_v^{\times p}$ for each $v \in S$. Then $E^{\times p} \subset Z_S$. Let $B = Z_S/E^{\times p}$ be the quotient $\mathbb{F}_p[H]$-module.

We say $\overline{\rho}$ is **tame** if the size of the image of $\overline{\rho}$ is prime to p. We say $\overline{\rho}$ is **regular** if it is absolutely irreducible, odd, and V and B are prime-to-adjoint.

Example. If $\overline{\rho}$ is tame and regular, then $R(\overline{\rho}) \cong \mathbb{Z}_p[\![T_1, T_2, T_3]\!]$.

For a concrete example, let E be the splitting field of $X^3 - X - 1$ over $\mathbb{Q}$. Then E is unramified away from 23 and ∞. The Galois group $\mathrm{Gal}(E/\mathbb{Q}) \cong S_3$. Since S_3 has a faithful representation in $\mathrm{GL}_2(\mathbb{F}_{23})$, we get an absolutely irreducible residual representation

$$\overline{\rho} : G_{\mathbb{Q},\{23,\infty\}} \to \mathrm{GL}_2(\mathbb{F}_{23}).$$

Its universal deformation ring is isomorphic to $\mathbb{Z}_{23}[\![T_1, T_2, T_3]\!]$.

If $\overline{\rho}$ is irregular, the situation is a bit more complicated.

Consider the residual representation

$$\overline{\rho} : G_{\mathbb{Q},\{3,7,\infty\}} \to \mathrm{GL}_2(\mathbb{F}_3)$$

coming from the elliptic curve $X_0(49)$. Then the universal deformation ring is isomorphic to $\mathbb{Z}_3[\![T_1, T_2, T_3, T_4]\!]/((1+T_4)^3 - 1)$.

Example. Let D be an integer congruent to $-1 \pmod 3$ and also $\pm$ a power of 2, and let $E/\mathbb{Q}$ be the elliptic curve defined by $y^2 = x(x^2 - 4Dx + 2D^2)$. Then E has complex multiplication by $\mathbb{Q}(\sqrt{-2})$. Let $S = \{2, 3\}$, and let $\overline{\rho} : G_{\mathbb{Q},S} \to \mathrm{GL}_2(\mathbb{F}_3)$ be the representation associated to E. Then $R(\overline{\rho}) \cong \mathbb{Z}_3[\![T_1, T_2, T_3, T_4, T_5]\!]/I$, where I is an ideal that takes a while to define.

We can give bounds on the number of (profinite) generators and relations it takes to present a deformation ring. Samit showed the following in his lecture in the fall:

Theorem 3. *Let K be a p-adic field, and let $G = G_K$ (for example). Let $r = \dim Z^1(G, \mathrm{Ad}(\overline{\rho}))$ and $s = \dim H^2(G, \mathrm{Ad}(\overline{\rho}))$. Then $R^{\square}(\overline{\rho})$ exists, and can be presented as*

$$R^{\square}(\overline{\rho}) \cong \mathfrak{o}_K[\![T_1, \dots, T_r]\!]/(f_1, \dots, f_s).$$

In the unframed case, we have $\dim R(\overline{\rho}) \geq 2 - \tilde{\chi}(G, \mathrm{Ad}(\overline{\rho}))$.

We also have

$$\dim(R/pR) \geq \dim H^1(G, \mathrm{Ad}(\overline{\rho})) - \dim H^2(G, \mathrm{Ad}(\overline{\rho})),$$

where the left side is the Krull dimension.

5. Characteristic zero points of deformation rings

Let S be a finite set of places of $\mathbb{Q}$. Fix an absolutely irreducible residual representation $\overline{\rho} : G_{\mathbb{Q},S} \to \mathrm{GL}_N(k)$, and let $\rho : G_{\mathbb{Q},S} \to \mathrm{GL}_N(R)$ be its universal deformation. We saw earlier that R looks something like $\mathbb{Z}_p[\![x]\!]$, but $R[1/p]$ is still far from being a local ring: in the case of $R = \mathbb{Z}_p[\![x]\!]$, $R[1/p] = \mathbb{Z}_p[\![x]\!][1/p] \subsetneq \mathbb{Q}_p[\![x]\!]$, since the power series on the left side need to have as denominators bounded powers of p. There are many $\mathbb{Q}_p$-algebra homomorphisms $\mathbb{Z}_p[\![x]\!][1/p] \twoheadrightarrow \mathfrak{o}_K[1/p]$ for various finite extensions $K/\mathbb{Q}_p$, where the first map sends x to a uniformizer of K. Thus, $R[1/p]$ has lots of maximal ideals, in this case. Something similar holds for general universal deformation rings. The maximal ideals of $R[1/p]$ correspond to deformations of $\overline{\rho}$ landing in finite extensions of $\mathbb{Q}_p$.

If W is a complete DVR, and R is a quotient of a power series ring in several variables over W, and ϖ is a uniformizer of W, then $R[1/\varpi]/\mathfrak{m}$ is finite over $W[1/\varpi]$ for any $\mathfrak{m} \in \mathrm{MaxSpec}(R[1/\varpi])$.

We'd like to understand what $R[1/p]$ looks like. Let $R = W[\![x_1, \dots, x_n]\!]/(f_1, \dots, f_m)$. Then, for any finite extension A of W,

$$\mathrm{Hom}(R, A) = \mathrm{Hom}(R[1/p], A[1/p] = \mathrm{Frac}(A)),$$

where the first Hom is in the category of local W-algebras, and the second is in the category of $\mathrm{Frac}(W)$-algebras.

Proposition 4. *If K'/K is a finite extension, then any K-algebra map $R[1/p] \to K'$ is given by sending the X_i's to various $x_i \in \mathfrak{m}_{K'} \subset \mathfrak{o}_{K'} \subset K'$. Hence the image of R lands in the valuation ring.*

Fix such a map $x : R[1/p] \twoheadrightarrow K'$. Let

$$\rho_x : G \xrightarrow{\rho} \mathrm{GL}_N(R) \to \mathrm{GL}_N(R[1/p]) \to \mathrm{GL}_N(K')$$

be the induced representation. We'd like to understand $\dim R[1/p]_{\mathfrak{m}_x} = \dim R[1/p]^{\wedge}_{\mathfrak{m}_x}$.

Theorem 5. *Let $\rho_x^{\mathrm{univ}} : G \to \mathrm{GL}_N(R[1/p]^\wedge_{\mathfrak{m}_x})$ be induced from ρ by the natural map $R \to R[1/p]^\wedge_{\mathfrak{m}_x}$. Then the diagram*

$$\begin{array}{ccc} G & \xrightarrow{\rho_x^{\mathrm{univ}}} & \mathrm{GL}_N(R[1/p]^\wedge_{\mathfrak{m}_x}) \\ & \searrow^{\rho_x} & \downarrow \\ & & \mathrm{GL}_N(K') \end{array}$$

commutes, and ρ_x^{univ} is universal for continuous deformations of ρ_x.

This theorem is relevant for (at least) two reasons:

(1) We have $R[1/p] \cong K'[\![T_1, \dots, T_n]\!]$ if and only if each $R[1/p]^\wedge_{\mathfrak{m}_x}$ is regular, if and only if the deformation functor is formally smooth, if and only if $H^2(G, \mathrm{Ad}(\rho_x)) = 0$.
(2) We have $(\mathfrak{m}_x/\mathfrak{m}_x^2)^\vee \cong H^1_{\mathrm{cts}}(G, \mathrm{Ad}(\rho_x))$ by the continuity condition on the deformations in the theorem.

Let $\rho : G = G_{\mathbb{Q},S} \to \mathrm{GL}_n(K)$ for some p-adic field K be a representation with absolutely irreducible reduction. Then $H^1(G, \mathrm{Ad}(\overline{\rho}))$ is also equal to the tangent space of the deformation ring $R(\overline{\rho})$ at the reduction of the closed point of $R[1/p]$ corresponding to ρ. Hence, the completion of $R[1/p]$ at that point (with scalars extended to K, if necessary), is the deformation ring for ρ.

6. Wiles product formula

Recall the definition of unramified cohomology. Let K be a p-adic field. If M is a K-module, then the unramified cohomology is

$$H^i_{nr}(K, M) = H^i(\mathrm{Gal}(K^{nr}/K), M^{I_K}).$$

If K is a global field, then for every place v of K, we have a map $G_{K_v} \hookrightarrow G_K$, and if M is a G_K-module, we have a restriction map $H^i(G_K, M) \to H^i(G_{K_v}, M)$ for each i. Let $\mathcal{L} = (L_v)$ be a collection of subgroups $L_v \subset H^1(G_{K_v}, M)$ so that $L_v = H^1_{nr}(G_{K_v}, M)$ for almost all v. Let

$$H^1_{\mathcal{L}}(G_K, M) = \{c \in H^1(G_K, M) \mid \mathrm{res}_v(c) \in L_v \text{ for all } v\}.$$

Let $\mathcal{L}^D = (L_v^D)$, where L_v^D is the annihilator of L_v under the Tate local pairing. We call the L_v the local conditions.

Theorem 6 (Wiles Product Formula)**.** *Suppose M is a finite G_K-module, $M' = \mathrm{Hom}(M, \mu)$, and $\mathcal{L}$ is a family of local conditions. Then*

$$\frac{\#H^1_{\mathcal{L}}(K, M)}{\#H^1_{\mathcal{L}^D}(K, M')} = \frac{\#H^0(K, M)}{\#H^0(K, M')} \prod_v \frac{\#L_v}{\#H^0(K_v, M)}.$$

We'll soon get to situations in which the denominator on the left-hand side of the Wiles Product Formula is 1. Thus, we'll have a formula for the size of the global H^1 in terms of sizes of H^0 as well as local terms. The local terms we'll be able to compute by studying various local deformation rings.

7. Example

Note: I don't really understand this example. I more or less copied Brian's email, but I'm including it for other people, who can probably understand it.

Let K be a p-adic field (p odd), and let $\omega : G_K \to k^\times$ be the mod-p cyclotomic character. Let $\rho : G_K \to \mathrm{GL}(V)$ be a residual representation, where V is a 2-dimensional k-vector space. Suppose the inertia group I_K acts nontrivially on V. Then let D be the subspace fixed by I_K. With respect to a suitable basis, then, we have

$$\rho = \begin{pmatrix} \theta_2 & * \\ 0 & \theta_1 \end{pmatrix},$$

where θ_1 and θ_2 are characters. We can write θ_1 and θ_2 uniquely in the form

$$\theta_1 = \omega^\alpha \varepsilon_1, \qquad \theta_2 = \omega^\beta \varepsilon_2,$$

where $\alpha, \beta \in \mathbb{Z}/(p-1)\mathbb{Z}$, and ε_1 and ε_2 are unramified characters $G_K \to k$. Hence the restriction of ρ to I_K is

$$\rho\,|_{I_K} = \begin{pmatrix} \omega^\beta & * \\ 0 & \omega^\alpha \end{pmatrix}.$$

We can normalize the exponents so that $0 \le \alpha \le p-2$ and $1 \le \beta \le p-1$.

Definition 7. If $\beta \neq \alpha + 1$, we say that ρ is **peu ramifié**. If $\beta = \alpha + 1$, we say that ρ is **très ramifié**.

Let $\mathcal{L}$ be a line in $H^1(K, \omega)$ not in the peu ramifié hyperplane. Let $\mathcal{H}$ be its orthogonal hyperplane in $H^1(K, k)$ with respect to the Tate pairing. Thus, $\mathcal{H}$ is a hyperplane not containing the unramified line. We wish to find a ramified character $\psi : G_K \to \Lambda^\times$ of finite order on I_K that lifts the trivial residual character and so that the image of $H^1(K, \varepsilon\psi) \to H^1(K, \omega)$ contains $\mathcal{L}$, where $\varepsilon : G_K \to \mathbb{Z}_p^\times$ is the p-adic cyclotomic character. Varying through nonzero points of $\mathcal{L}$ gives us a collection of non-isomorphic non-split extensions of a common reducible but indecomposable G_K-module, so gluing them together gives us a lifting result for 2-dimensional Galois representations.

ψ allows for a lift of $\mathcal{L}$ if and only if $\mathcal{L}$ is in the kernel of the connecting map to $H^2(K, \varepsilon\psi)$. Let F be the fraction field of Λ, and let ϖ be a uniformizer. This happens if and only if connecting map

$$H^0(K, (F/\Lambda)(\psi^{-1})) \to H^1(K, k)$$

attached to the sequence

$$0 \to k \to (F/\Lambda)(\psi^{-1}) \xrightarrow{\varpi} (F/\Lambda)(\psi) \to 0$$

has image contained in $\mathcal{H}$. So, let's figure out exactly what the connecting map does in order to see what it means on the ramified character ψ^{-1} of finite order on I_K that the image is contained in some hyperplane not containing the unramified line.

If $x = u\varpi^{-n}$ for $n > 0$ and $u \in \Lambda^*$, we have $x \in H^0(K, (F/\Lambda)(\psi^{-1}))$ if and only if $\psi^{-1} \equiv 1 \pmod{\varpi^n}$. Since ψ (and hence ψ^{-1}) is nontrivial, this only works for finitely many values of n, but including $n = 1$. The image of x under this connecting map is the k-torsor of points

$$(u\varpi^{-n})(\varpi^{-1} + \Lambda) \mod \Lambda,$$

and the corresponding character $G_K \to k = \varpi^{-1}\Lambda/\Lambda$ is

$$\phi_n : g \mapsto (\varpi^{-1} + \Lambda)((\psi^{-1}(g) - 1)\varpi^{-n}) \mod \Lambda = (\psi^{-1}(g) - 1)\varpi^{-n-1} \mod \Lambda.$$

Note that $(\psi^{-1}(g)-1)\varpi^{-n} \in \Lambda$, and if n is not maximal with respect to this property, then $\phi_n = 0$. If we write

$$\psi^{-1} = 1 + \varpi^n\chi$$

and $\chi_0 = \chi \mod \varpi$, then χ_0 is a nontrivial character $G_K \to k$, and $1+\varpi^n\chi$ restricted to I_K is valued in the p^{th} power roots of unity in $\Lambda^\times$. The condition on ψ^{-1} is that χ_0 is contained in $\mathcal{H} \subset H^1(K, k)$.

More concretely, this is equivalent to the following. Let $\mathcal{H}$ be a hyperplane in $H^1(K, k)$ not containing the unramified line. We seek a continuous character $\xi : G_K \to 1 + \mathfrak{m}_\Lambda$ with finite order on I_K and conductor $n > 0$ so that the nonzero additive character $(\xi - 1)\varpi^{-n} : G_K \to k$ lies in $\mathcal{H}$. This character *must* be ramified. We could have replaced ϖ^n with $u\varpi^n$ for any $u \in \Lambda^\times$.

In order to make the construction, we need to start with a Λ containing a primitive p^{th} root of unity ζ. Let $n = e/(p-1)$, where $e = e(\Lambda)$, so that $\zeta - 1 = \varpi^n$. Fix a nontrivial character $\xi : \mathfrak{o}_K^\times \to \mu_p$, and extend it to an order p character ϕ on G_K by class field theory, so that $\phi \equiv 1 \mod \mathfrak{m}_A^n$. The function

$$\chi = \frac{\phi - 1}{\zeta - 1} : G_K \to k$$

is an additive character that is not identically zero, and it is ramified: there is an element $\tau \in I_K$ taken to ζ^{-1}.

If $\chi \in \mathcal{H}$, we're done. If not, then $\mathcal{H}$ is a hyperplane not containing the unramified line, so any element not in $\mathcal{H}$ can be translated by a unique unramified k-valued character so that it does lie in $\mathcal{H}$. For a unit $u_0 \equiv 1 \mod \mathfrak{m}_A^{e/(p-1)}$, twisting by the unramified character taking Frob to u_0 has the effect of adding to χ the unramified character taking Frob to $(u_0 - 1)/(\zeta - 1)$. Varying u_0, this sweeps through all the unramified characters $G_K \to k$, so we can hit the one we need to land in $\mathcal{H}$.

Calculating deformation rings

Rebecca Bellovin

1 Introduction

We are interested in computing local deformation rings away from p. That is, if L is a finite extension of $\mathbf{Q}_\ell$ and V is a 2-dimensional representation of G_L over $\mathbf{F}$, where $\mathbf{F}$ is a finite extension of $\mathbf{F}_p$, $\ell \neq p$, we wish to study the deformation rings $R_V^{\square}$ and $R_V^{\psi,\square}$. Here $\psi : G_L \to \mathcal{O}^\times$ is a continuous unramified character, $\mathcal{O}$ is the ring of integers of a finite extension E of $\mathbf{Q}_p$ which has residue field $\mathbf{F}$, and $R_{V_{\mathbf{F}}}^{\psi,\square}$ is the quotient of $R_{V_{\mathbf{F}}}^{\square}$ corresponding to deformations with determinant $\psi\chi$, where $\chi : G_L \to \mathbf{Z}_p^\times$ is the cyclotomic character.

Note that $R_{V_{\mathbf{F}}}^{\psi,\square}$ exists: There is a natural determinant map from the universal 2-dimensional (framed) representation to the universal 1-dimensional (framed) representation, and we take the fiber over the closed point corresponding to $\chi\psi$.

We define the following two deformation problems:

- $D_V^{ur,\psi,\square}$ is the deformation functor which spits out unramified framed deformations with determinant $\psi\chi$
- $L_V^{\chi,\square}$ is the deformation functor which spits out pairs (V_A, L_A) of framed deformations with determinant χ, together with a G_L-stable A-line with G_L acting via χ on L_A. That is, L_A is a projective rank 1 A-module such that V_A/L_A is a projective A-module with trivial G_L-action.

Most of this talk will be about the structure of the ring representing the second functor.

2 Lies I will tell, and auxiliary categories of rings

The minor lie I will tell is that I will entirely suppress the language of categories fibered in groupoids, and pretend we are working with functors. This will allow me to avoid 2-categorical language. But to make what I say literally true, one has to handle non-trivial isomorphisms of deformations via the language of groupoids.

The more major lie I will tell is that after I finish this section, I will try to avoid talking about the various categories of algebras that are involved.

The basic set-up is representing certain deformations of a fixed residual representation (in characteristic p). The deformations are a priori to finite local artinian rings with fixed residue field. But we want to be able to take generic fibers of our representing objects in a sensible way, so we need techniques for passing to characteristic 0 points.

To do this, we need a variety of confusing auxiliary categories of algebras. To demonstrate, let $E/\mathbf{Q}_p$ be a finite extension with residue field containing $\mathbf{F}$, let $\mathcal{O} \subset \mathcal{O}_E$ be a discrete valuation ring finite over $W(\mathbf{F})$, and let D be a deformation functor on the category $\mathfrak{AA}_{\mathcal{O}}$ of finite local artinian $\mathcal{O}$-algebras with residue field $\mathcal{O}/\mathfrak{m}_{\mathcal{O}}$, and let $E/\mathbf{Q}_p$ be a finite extension with residue field containing $\mathbf{F}$. We will be interested in the category $\mathfrak{AR}_E$ of finite local $W(\mathbf{F})[1/p]$-algebras with residue field E. We also introduce the following categories:

$\widehat{\mathfrak{AR}}_{\mathcal{O}}$: $\widehat{\mathfrak{AR}}_{\mathcal{O}}$ is the category of complete local noetherian $\mathcal{O}$-algebras with residue field $\mathcal{O}/\mathfrak{m}_{\mathcal{O}}$.

$\widehat{\mathfrak{AR}}_{\mathcal{O},(\mathcal{O}_E)}$: $\widehat{\mathfrak{AR}}_{\mathcal{O},(\mathcal{O}_E)}$ is the category of $\mathcal{O}$-algebras A in $\widehat{\mathfrak{AR}}_{\mathcal{O}}$ equipped with a map of $\mathcal{O}$-algebras $A \to \mathcal{O}_E$.

Int_B: Given $B \in \mathfrak{AR}_E$, Int_B is the category of finite $\mathcal{O}_E$-subalgebras $A \subset B$ such that $A \otimes_{\mathcal{O}_E} E = B$.

Note that Int_B is a subcategory of $\widehat{\mathfrak{AR}}_{\mathcal{O},(\mathcal{O}_E)}$ (A obviously has a map to E, and by finiteness or the same sort of arguments as in Brian's talk, it actually lands in $\mathcal{O}_E$), and there is a natural functor $\widehat{\mathfrak{AR}}_{\mathcal{O},(\mathcal{O}_E)} \to \widehat{\mathfrak{AR}}_{\mathcal{O}}$.

Also note that we can canonically extend D to a groupoid on $\widehat{\mathfrak{AR}}_{\mathcal{O}}$, by setting $D(\varprojlim R/\mathfrak{m}_R^{n+1}) = \varprojlim D(R/\mathfrak{m}_R^{n+1})$.

Now fix some $\xi \in D(\mathcal{O}_E)$, which makes sense by the preceding comment. We define a groupoid $D_{(\xi)}$ on $\widehat{\mathfrak{AR}}_{\mathcal{O},(\mathcal{O}_E)}$ by setting $D_{(\xi)}$ to be the fiber over ξ. More precisely, $D_{(\xi)}(A)$ consists of objects of $D(A)$ together with morphisms (in D) covering the given map $A \to \mathcal{O}_E$.

Finally, we can extend $D_{(\xi)}$ to $\mathfrak{AR}_E$. We note that $B \in \mathfrak{AR}_E$ can be exhausted by objects in Int_B, so we set $D_{(\xi)}(B) = \varinjlim_{A \in \mathrm{Int}\, B} D_{(\xi)}(A)$.

Now Kisin proves two crucial lemmas about these groupoids (which he calls a lemma and a proposition). The first tells us how to get universal deformation rings for the groupoids on $\mathfrak{AR}_E$, and the second relates those groupoids to the ones we would naively expect (for some deformation problems we already care about):

Lemma 2.1. *If D is pro-represented by a complete local $\mathcal{O}$-algebra R, then $D_{(\xi)}$ is pro-represented (on $\mathfrak{AR}_E$) by the complete local $\mathcal{O}[1/p]$-algebra $\hat{R}_\xi$ obtained by completing $R \otimes_{\mathcal{O}} E$ along the kernel I_ξ of the map $R \otimes_{\mathcal{O}} E \to E$ induced by ξ.*

Lemma 2.2. *Fix a residual representation V over $\mathbf{F}$, and carry out the above program for D_V and $D_V^\square$. Then there are natural isomorphisms of groupoids*

$$D_{V,(\xi)} \tilde{\to} D_{V_\xi} \text{ and } D^\square_{V,(\xi)} \tilde{\to} D^\square_{V_\xi}$$

3 Main result

The main result we will prove is the following:

Theorem 3.1. *Let V be any 2-dimensional representation of G_L (over $\mathbf{F}$). Fix a continuous unramified character $\psi : G_L \to \mathcal{O}^\times$ and consider $R_V^{\psi,\square}$, the quotient of $R_V^\square$ corresponding to deformations of V with determinant $\psi\chi$. Then $\operatorname{Spec} R_V^{\psi,\square}[1/p]$ is 3-dimensional, and it is the scheme-theoretic union of formally smooth components.*

There are several claims implicit in this theorem, namely the existence, smoothness, connectedness, and dimension of $R_V^{\mathrm{ur},\psi,\square}$ and $R_V^{\chi\gamma,\gamma,\square}$, as well as the connectedness of $R_V^{\psi,\square}$. We assume these for the moment and go on with the proof.

Proof. Let E'/E be a finite extension, let $x : R_V^{\psi,\square}[1/p] \to E'$ be a point of $\operatorname{Spec} R_V^{\psi,\square}[1/p]$ with residue field E' (so that it is actually an E'-point), and let V_x be the induced representation with coefficients in E'. We know (from Brian's talk on characteristic 0 points of deformation rings) that the completion of $R_V^{\psi,\square}[1/p]$ at the maximal ideal $\mathfrak{m}_x = \ker x$ represents deformations of V_x. The tangent space at x is $H^1(G_L, \operatorname{ad}^0 V_x)$. Obstructions to deforming representations live in H^2 groups, so $R_V^{\psi,\square}[1/p]$ at x will be formally smooth at any point x where $H^2(G_L, \operatorname{ad}^0 V_x)$ vanishes.

Given any framed deformation problem $D^\square$ (with coefficients in some unspecified field H), there is a natural morphism $D^\square \to D$ to the unframed problem given by "forgetting the basis". This morphism is formally smooth in the sense that artinian points of D can be lifted.

Furthermore, the fibers of the morphism of tangent spaces $D^\square(H[\varepsilon]) \to D(H[\varepsilon])$ are principal homogeneous spaces under $\operatorname{ad} / \operatorname{ad}^{G_L}$. Specifically, given a residual representation V_H and a choice of (unframed) deformation $V_{H[\varepsilon]}$, $\ker(\mathrm{GL}_2(H[\varepsilon]) \to \mathrm{GL}_2(H)) = 1 + \varepsilon M_2(H[\varepsilon]) \cong \operatorname{End}_H V_H$ acts (via conjugation) on the fiber over $V_{H[\varepsilon]}$. Then it is easy to check that $1 + \varepsilon M$ acts trivially on the fiber if and only if M is in $\operatorname{ad}^0 V_H$.

Counting dimensions,

$$\dim_F D^\square(\mathbf{F}[\varepsilon]) = \dim_F D(\mathbf{F}[\varepsilon]) + \dim_F \operatorname{ad} - \dim_F H^0(G_L, \operatorname{ad}) \tag{3.1}$$

Using this formula, we see that the tangent space to $\operatorname{Spec} R_V^{\psi,\square}[1/p]$ at x has E'-dimension

$$\begin{aligned}
&\dim_{E'} H^1(G_L, \operatorname{ad}^0 V_x) + \dim_{E'} \operatorname{ad} V_x - \dim_{E'} H^0(G_L, \operatorname{ad} V_x) \\
&\quad = \dim_{E'} H^1(G_L, \operatorname{ad}^0 V_x) + \dim_{E'} \operatorname{ad} V_x - (\dim_{E'} H^0(G_L, \operatorname{ad}^0 V_x) - 1) \\
&\quad = -\big(\dim_{E'} H^2(G_L, \operatorname{ad}^0 V_x) - \dim_{E'} H^1(G_L, \operatorname{ad}^0 V_x) + \dim_{E'} H^0(G_L, \operatorname{ad}^0 V_x)\big) \\
&\qquad + \dim_{E'} H^2(G_L, \operatorname{ad}^0 V_x) + \dim_{E'} \operatorname{ad} V_x - 1 \\
&\quad = \dim_{E'} H^2(G_L, \operatorname{ad}^0 V_x) + 3
\end{aligned}$$

the last step following by the Euler characteristic formula for p-adic coefficients. Thus, if $H^2(G_L, \operatorname{ad}^0 V_x) = 0$, x will be a formally smooth point of $\operatorname{Spec} R_V^{\psi,\square}[1/p]$ with a 3-dimensional tangent space.

Now suppose $H^2(G_L, \operatorname{ad}^0 V_x) \neq 0$. By the p-adic version of Tate local duality, $\dim_{E'} H^2(G_L, \operatorname{ad}^0 V_x) = \dim_{E'} H^0(G_L, (\operatorname{ad}^0 V_x)^*)$, which is $\dim_{E'} H^0(G_L, \operatorname{ad}^0 V_x(1))$

(because ad^0 is self-dual). Now we have the split exact sequence of G_L-modules

$$0 \to \mathrm{ad}^0 V_x(1) \to \mathrm{ad}\, V_x(1) \to E'(1) \to 0$$

which gives us an exact sequence in cohomology:

$$0 \to H^0(G_L, \mathrm{ad}^0 V_x(1)) \to H^0(G_L, \mathrm{ad}\, V_x(1)) \to H^0(G_L, E'(1))$$

But $H^0(G_L, E'(1)) = 0$ so

$$H^0(G_L, \mathrm{ad}^0 V_x(1) = H^0(G_L, \mathrm{ad}\, V_x(1)) = H^0(G_L, \mathrm{Hom}(V_x, V_x(1)))$$

In particular, if $H^2(G_L, \mathrm{ad}^0 V_x) \neq 0$, there is a non-zero homomorphism (of G_L-modules) $V_x \to V_x(1)$. It has 1-dimensional (G_L-stable) image and kernel, so there is some character γ such that $0 \to \gamma \to V_x \to \gamma(1) \to 0$ is exact. But such extensions are classified by $H^1(G_L, E'(-1))$, which is 0: the Euler characteristic formula says that $\dim_{E'} H^0(G_L, E'(-1)) - \dim_{E'} H^1(G_L, E'(-1)) + \dim_{E'} H^2(G_L, E'(-1)) = 0$, but $H^0(G_L, E'(-1))$ is clearly zero, and $H^2(G_L, E'(-1))$ is dual to $H^0(G_L, E'(2))$, which is zero, so $H^1(G_L, E'(-1))$ is zero as well. So this extension splits.

We have shown that if $H^2(G_L, \mathrm{ad}^0 V_x) \neq 0$, then $V_x = \gamma \oplus \gamma\chi$ for some character $\gamma : G_L \to E'^{\times}$. If γ is unramified, then this implies that x is in the image of both $R_V^{\mathrm{ur},\gamma^2,\square}$ and $R_V^{\chi\gamma,\gamma,\square}$.

So the only singular points of $\mathrm{Spec}\, R_V^{\psi,\square}[1/p]$ lie in the intersection of two formally smooth components.

□

The definition of formal smoothness requires us to be able to lift through *any* square-zero thickening, but we only looked at what happens at artinian points of $\mathrm{Spec}\, R_V^{\psi,\square}[1/p]$; the commutative algebra necessary to justify this is discussed in Brian's notes on $\ell = p$.

4 Unramified deformations

We've seen previously that for the unframed case, the tangent space at x for unramified deformations with fixed determinant is $H^1(G_L/I_L, (\mathrm{ad}^0 V_x)^{I_L})$,

and the obstruction space should be $H^2(G_L/I_L, (\mathrm{ad}^0 V_x)^{I_L}) = 0$. We have the exact sequence

$$0 \to (\mathrm{ad}^0 V_x)^{G_L} \to (\mathrm{ad}^0 V_x)^{I_L} \xrightarrow{\mathrm{Frob}-\mathrm{id}} (\mathrm{ad}^0 V_x)^{I_L} \to (\mathrm{ad}^0 V_x)^{I_L}/(\mathrm{Frob}-\mathrm{id})(\mathrm{ad}^0 V_x)^{I_L} \to 0$$

This implies that $\dim_{E'} H^0(G_L, \mathrm{ad}^0 V_x) = \dim_{E'} H^1(G_L/I_L, (\mathrm{ad}^0 V_x)^{I_L})$. And since the tangent space for the framed case has dimension $\dim_{E'} H^1(G_L/I_L, (\mathrm{ad}^0 V_x)^{I_L}) + \dim_{E'} \mathrm{ad}^0 V_x - \dim_{E'} H^0(G_L, \mathrm{ad}^0 V_x)$ by the discussion in the previous section, this implies that the tangent space of $R_V^{ur,\psi,\square}$ has dimension $\dim_{E'} \mathrm{ad}^0 V_x = 3$.

So granting existence, $R_V^{ur,\psi,\square}$ is formally smooth and 3-dimensional.

5 $R^{\chi\gamma,\gamma,\square}$

We begin this section with a general lemma.

Lemma 5.1. *Let $\mathcal{O}$ be a local $W(k)$-algebra with residue field k, with K the fraction field of $W(k)$, and let X be a proper residually reduced $\mathcal{O}$-scheme. Then the components of the fiber of X over the closed point of $\mathcal{O}$ are in bijection with the components of $X[1/p]$.*

Proof. Consider a connected component of $X[1/p] = X \otimes_{W(k)} K$ and let e be the idempotent which is 1 on this component and 0 on the others. Then if ϖ is a uniformizer of $W(k)$, there is some n such that $\varpi^n e$ extends to a global section of X. But $(\varpi^n e)^2 = \varpi^n(\varpi^n e)$, so if $n > 0$, as a function on the special fiber $X \otimes_{\mathcal{O}} k$, $\varpi^n e$ is nilpotent. This contradicts our reducedness hypothesis, so $n = 0$ and e is already a global section of X.

So we know that the components of $X \otimes_{W(k)} K$ are in bijection with the components of X itself. But if $X^\wedge$ is the completion of X along its special fiber, the components of the special fiber $X \otimes_{\mathcal{O}} k$ are in bijection with the components of $X^\wedge$ (because they have the same underlying topological space), and formal GAGA implies that the components of $X^\wedge$ are in bijection with the components of X (X is proper over $\mathcal{O}$, so we can apply formal GAGA to see that the global idempotent functions on X and $X^\wedge$ are in bijection). $\square$

5.1 Representability

Proposition 5.2. *The morphism* $|L_V^{\chi,\square}| \to |D_V^{\chi,\square}|$ *is represented by a projective morphism* $\Theta_V : \mathcal{L}_V^{\chi,\square} \to R_V^{\chi,\square}$.

Proof. Given an A-point of $R_V^{\chi,\square}$, the A-points of $\mathcal{L}_V^{\chi,\square}$ should be certain line bundles on $\operatorname{Spec} A$, so we will cut $\mathcal{L}_V^{\chi,\square}$ out of $\mathbf{P}^1_{R_V^{\chi,\square}}$.

Consider $\mathbf{P}$, the projectivization of the universal rank 2 $R_V^{\chi,\square}$-module. That is, if V_R is the universal rank 2 $R_V^{\chi,\square}$-module (equipped with a representation of G_L), then $\mathbf{P} := \operatorname{Proj} \operatorname{Sym} V_R \cong \operatorname{Proj} R_V^{\chi,\square}[x_0, x_1]$.

If A is an an $R_V^{\chi,\square}$-algebra with residue field $\mathbf{F}$, a morphism $\operatorname{Spec} A \to \mathbf{P}$ (over $R_V^{\chi,\square}$) is the same as a surjection (of sheaves) $A^2 \to \mathcal{L} \to 0$.

Given a morphism $f : \operatorname{Spec} A \to \mathbf{P}$, there is a natural G_L-action on the quotient $\mathcal{L}$ if and only if $g^* f = f$ for all $g \in G_L$. The g^*-fixed locus of $\mathbf{P}$ is H_g defined by the Cartesian square

$$\begin{array}{ccc} H_g & \longrightarrow & \mathbf{P} \\ \downarrow & & \downarrow{\scriptstyle (\mathrm{id}, g^*)} \\ \mathbf{P} & \xrightarrow{\Delta} & \mathbf{P} \times_{R_V^{\chi,\square}} \mathbf{P} \end{array}$$

Since $\mathbf{P}$ is separated, H_g is a closed subscheme of $\mathbf{P}$. Thus, the intersection $H := \cap_{g \in G} H_g$ is a closed subscheme of $\mathbf{P}$ parametrizing G_L-equivariant quotients $A^2 \to \mathcal{L} \to 0$.

Now if A is a complete local $W(\mathbf{F})$-algebra, there is a natural map from H to the universal deformation of the residually trivial 1-dimensional representation, given (in the language of the functor of points) by sending $A^2 \to \mathcal{L} \to 0$ to $\mathcal{L}$. Then we can take the fiber over the (closed) point corresponding to the trivial representation to get a closed subscheme of $\mathbf{P}$ representing $L_V^{\chi,\square}$ on $\mathfrak{AR}_{W(\mathbf{F})}$.

Now take limits to get representability of $L_V^{\chi,\square}$ on $\mathfrak{Aug}_{W(\mathbf{F})}$. □

5.2 Smoothness and connectedness

Next we want to study smoothness and connectedness.

Proposition 5.3. *$\mathcal{L}_V^{\chi,\square}$ is formally smooth over $W(\mathbf{F})$. Furthermore, the $W(\mathbf{F})[1/p]$-scheme $\mathcal{L}_V^{\chi,\square} \otimes_{W(\mathbf{F})} W(\mathbf{F})[1/p]$ is connected.*

Proof. First, we will show that for any finite group M of p-power order, the natural map $H^1(G_L, \mathbf{Z}_p(1)) \otimes_{\mathbf{Z}_p} M \to H^1(G_L, \mathbf{Z}_p(1) \otimes_{\mathbf{Z}_p} M)$ is an isomorphism. It suffices to consider the case $M = \mathbf{Z}/p^n\mathbf{Z}$. In that case, we have the exact sequence

$$0 \to \mathbf{Z}_p(1) \xrightarrow{\cdot p^n} \mathbf{Z}_p(1) \to M \to 0$$

Then the long exact sequence in group cohomology shows that

$$0 \to H^1(G_L, \mathbf{Z}_p(1))/p^n H^1(G_L, \mathbf{Z}_p(1)) \to H^1(G_L, M) \to H^2(G_L, \mathbf{Z}_p(1))[p^n]$$

is exact. The middle arrow is the natural map we started with, so we wish to show that $H^2(G_L, \mathbf{Z}_p(1))[p^n]$ is 0. But by Tate local duality (as in Simon's talk), $H^2(G_L, \mathbf{Z}_p(1))$ is Pontryagin dual to $\mathbf{Q}_p/\mathbf{Z}_p$, so has no p^n-torsion.

Thus, for any artinian algebra A, the composition

$$\mathrm{Ext}^1_{\mathbf{Z}_p[G_L]}(\mathbf{Z}_p, \mathbf{Z}_p(1)) \otimes_{\mathbf{Z}_p} A \to H^1(G_L, \mathbf{Z}_p(1)) \otimes_{\mathbf{Z}_p} A \to H^1(G_L, \mathbf{Z}_p(1) \otimes_{\mathbf{Z}_p} A) \to \mathrm{Ext}^1_{\mathbf{Z}_p[G_L]}(A, A(1))$$

is an isomorphism.

To prove smoothness, it suffices to show that for any surjection of artinian rings $A \to A'$, the map $|L_V^{\chi,\square}|(A) \to |L_V^{\chi,\square}|(A')$ is a surjection. Now consider a pair $(V_{A'}, L_{A'})$ in $|L_V^{\chi,\square}|(A')$. It corresponds to an element of $\mathrm{Ext}^1_{\mathbf{Z}_p[G_L]}(A, A(1))$, so by the isomorphism we just proved, it corresponds to an element of $\mathrm{Ext}^1_{\mathbf{Z}_p[G_L]}(\mathbf{Z}_p, \mathbf{Z}_p(1)) \otimes_{\mathbf{Z}_p} A'$. But such an element clearly lifts to an element of $\mathrm{Ext}^1_{\mathbf{Z}_p[G_L]}(\mathbf{Z}_p, \mathbf{Z}_p(1)) \otimes_{\mathbf{Z}_p} A$, which is to say, an element of $|L_V^{\chi,\square}|(A)$.

Now we wish to prove connectedness after inverting p, and for this we use the lemma on connected components. Specifically, since $\mathcal{L}_V^{\chi,\square}$ is smooth, its special fiber $\mathcal{L}_V^{\chi,\square} \otimes_{W(\mathbf{F})} \mathbf{F}$ is reduced, so to show $\mathcal{L}_V^{\chi,\square}[1/p]$ is connected, it suffices to show that the special fiber $\mathcal{L} \otimes_{R_V^{\chi,\square}} \mathbf{F}$ is connected.

But the special fiber is simply the fiber over the residual representation. If $\mathbf{F} \cong \mathbf{F}(1)$ and the representation is split (i.e., the residual representation is trivial), any line in $\mathbf{F}^2$ is G_L-stable with G_L-acting by $\chi = \mathrm{id}$, so the fiber is a full $\mathbf{P}^1_{\mathbf{F}}$. Otherwise, there is at most one G_L-line with G_L acting via χ, and this is true for any A-point of the fiber, so it is either empty or it consists of a single reduced point. So the special fiber is connected. $\square$

The next proposition will show that $\mathcal{L}_V^{\chi,\square}[1/p] \to \operatorname{Spec} R_V^{\chi,\square}[1/p]$ is a monomorphism. More precisely, it shows that this morphism is injective on artinian points, but, as before, Brian's notes on $\ell = p$ explain why this is sufficient to let us conclude that it is actually a monomorphism.

Proposition 5.4. *Let $E/\mathbf{Q}_p$ be a finite extension, and let ξ refer to both an $\mathcal{O}_E$-valued point of $R_V^{\chi,\square}$ and an $\mathcal{O}_E$-valued point in the fiber of $\mathcal{L}_V^{\chi,\square}$ above it. Then the morphism of groupoids (functors) on $\mathfrak{AR}_E$ $L_{V_\xi}^{\chi,\square} \to D_{V_\xi}^{\chi,\square}$ is fully faithful. If the representation over E V_ξ corresponding to ξ is indecomposable, then this is an equivalence.*

Proof. Let B be an object of $\mathfrak{AR}_E$, and let V_B be an object of $D_{V_\xi}^{\chi,\square}(B)$. To prove the first assertion, we need to show that V_B admits at most one G_L-stable B-line $L_B \subset V_B$ such that G_L acts trivially on V_B/L_B. But $\operatorname{Hom}_{B[G_L]}(B(1), V_B/L_B) = \{0\}$ because the G_L-action on the target is trivial, so $\operatorname{Hom}_{B[G_L]}(B(1), V_B) = \operatorname{Hom}_{B[G_L]}(B(1), L_B)$ and L_B is unique.

Now suppose V_ξ is indecomposable; we wish to show that V_B actually does admit a suitable B-line. We will do this by showing that V_B is isomorphic to the trivial deformation $V_\xi \otimes_E B$. Note that by Tate local duality

$$\dim_E H^1(G_L, \operatorname{ad}^0 V_\xi) = \dim_E H^0(G_L, \operatorname{ad}^0 V_\xi) + \dim_E H^0(G_L, \operatorname{ad}^0 V_\xi(1)) = 0$$

the last equality following from indecomposability of V_ξ. The result then follows by induction on the length of B, since this calculation holds for any indecomposable extension of $A(1)$ by A. □

But since we have a proper monomorphism of schemes $\mathcal{L}_V^{\chi,\square}[1/p] \to \operatorname{Spec} R_V^{\chi,\square}[1/p]$, it is a closed immersion.

Now we can prove the following proposition and corollary.

Proposition 5.5. *Let $\operatorname{Spec} R_V^{\chi,1,\square}$ be the scheme-theoretic image of the morphism $\mathcal{L}_V^{\chi,\square} \to \operatorname{Spec} R_V^{\chi,\square}$. Then*

1. *$R_V^{\chi,1,\square}$ is a domain of dimension 4 and $R_V^{\chi,1,\square}$ is formally smooth over $W(\mathbf{F})$.*
2. *If $E/\mathbf{Q}_p$ is a finite extension, then a morphism $\xi : R_V^{\chi,\square} \to E$ factors through $R_V^{\chi,1,\square}$ if and only if the corresponding two-dimensional representation V_ξ is an extension of E by $E(1)$.*

Proof. Since $R_V^{\chi,1,\square}$ is smooth and connected, it is a domain. We will find its dimension via a tangent space calculation. Suppose V_ξ is indecomposable (which we may assume, since most points on $R_V^{\chi,1,\square}$ are indecomposable). Then the dimension of $R_V^{\chi,1,\square}[1/p]$ is

$$\begin{aligned}\dim_E |D_{V_\xi}^{\chi,\square}|(E[\varepsilon]) &= \dim_E |D_{V_\xi}^{\chi}|(E[\varepsilon]) + 4 - \dim_E(\operatorname{ad} V_\xi)^{G_L} \\ &= \dim_E H^1(G_L, \operatorname{ad}^0 V_\xi) + 3 = 3\end{aligned}$$

So $R_V^{\chi,1,\square}$ itself is 4-dimensional, and we have proven the first part.

The second part follows from the definition of $\mathcal{L}_V^{\chi,\square}$ and $R_V^{\chi,1,\square}$. □

Corollary 5.6. *Let $\mathcal{O}$ be the ring of integers in a finite extension of $W(\mathbf{F})[1/p]$, and $\gamma : G_L \to \mathcal{O}^\times$ a continuous unramified character. Write $R_{V,\mathcal{O}}^{\square} = R_V^{\square} \otimes_{W(\mathbf{F})} \mathcal{O}$. Then there exists a quotient $R_{V,\mathcal{O}}^{\chi\gamma,\gamma,\square}$ such that*

- *$R_{V,\mathcal{O}}^{\chi\gamma,\gamma,\square}$ is a domain of dimension 4 and $R_{V,\mathcal{O}}^{\chi\gamma,\gamma,\square}[1/p]$ is formally smooth over $\mathcal{O}$.*
- *If $E/\mathcal{O}[1/p]$ is a finite extension, then a map $\xi : R_{V,\mathcal{O}}^{\square} \to E$ factors through $R_{V,\mathcal{O}}^{\chi\gamma,\gamma,\square}$ if and only if V_ξ is an extension of γ by $\gamma(1)$.*

Proof. This basically follows because universal deformation rings behave reasonably well with respect to twisting by fixed characters, at least once the question makes sense.

More precisely, we may replace $\mathbf{F}$ by the residue field of $\mathcal{O}$ (corresponding to tensoring $R_V^{\square}$ with $\mathcal{O}$). Then twisting by γ^{-1} induces an isomorphism $R_{V,\mathcal{O}}^{\square} \tilde{\to} R_{V\times\gamma^{-1},\mathcal{O}}^{\square}$ (because twisting the residual representation by γ^{-1} doesn't change this deformation problem (except to multiply the determinant by γ^2), and the quotient $R_{V,\mathcal{O}}^{\chi\gamma,\gamma,\square}$ corresponds to $R_{V\otimes\gamma^{-1}}^{\chi,1,\square} \otimes_{W(\mathbf{F})} \mathcal{O}$ under this isomorphism. □

Lecture 21: Structure of ordinary-crystalline deformation ring for $\ell = p$

1. Basic problem

Let Λ be a complete discrete valuation ring with fraction field E of characteristic 0, maximal ideal $\mathfrak{m} = (\pi)$, and residue field k of characteristic $p > 0$; we will ultimately be interested in the case when k is finite (and in particular, perfect). Consider a complete local noetherian Λ-algebra R with residue field k, so

$$R = \Lambda[\![x_1, \ldots, x_m]\!]/(f_1, \ldots, f_s),$$

and suppose there is given a continuous representation

$$\rho : G_K \to \mathrm{GL}_n(R)$$

for a p-adic field K (i.e., K is a finite extension of $\mathbf{Q}_p$). Note that $R[1/p] = R[1/\pi] = R \otimes_\Lambda E$; we call this the "generic fiber" of R over Λ, but beware that as an E-algebra this is typically very far from being finitely generated. For shorthand, we write R_E to denote this generic fiber.

We are going to be interested in certain subsets of $\mathrm{MaxSpec}(R_E)$. Recall from the lecture in the fall on generic fibers of deformation rings that the maximal ideals of R_E are precisely the kernels of E-algebra homomorphisms $R_E \to E'$ into finite extensions E'/E, or equivalently Λ-algebra homomorphisms $R \to E'$, and that such maps are necessarily given by

$$h(x_1, \ldots, x_m) \mapsto h(a_1, \ldots, a_m)$$

for a_i in the maximal ideal of the valuation ring Λ' of E'. In more geometric terms, $\mathrm{MaxSpec}(R_E)$ is identified with the zero locus

$$\{(a_1, \ldots, a_m) \in \overline{E}^m \,|\, |a_i| < 1, f_j(a_1, \ldots, a_m) = 0 \text{ for all } j\}$$

taken up to the natural action by $\mathrm{Gal}(\overline{E}/E)$ on this locus.

Remark 1.1. Loosely speaking, we view R_E as an "algebraist's substitute" for working directly with the rigid-analytic space $\{f_1 = \cdots = f_s = 0\}$ inside of the open unit polydisk over E. There is a way to make this link more precise, by relating R_E to the algebra of bounded analytic functions on this analytic space, but we do not need such a result so we will pass over it in silence; nonetheless, trying to visualize $\mathrm{MaxSpec}(R_E)$ in terms of this analytic zero locus is a good source of intuition.

In the fall lecture on generic fibers of deformation rings, we recorded a few basic algebraic properties of R_E and we recall them now. First, R_E is noetherian and Jacobson; the latter means that every prime ideal is the intersection of the maximals over it, or equivalently the radical of any ideal is the intersection of the maximals over it. This ensures that focusing on MaxSpec does not lose a lot of information, much like algebras of finite type over a field (*the* classic example of a Jacobson ring). In contrast, a local ring of positive dimension (e.g., a discrete valuation ring, or R as above when not artinian!) is never Jacobson! An additional important property, already implicit in the preceding discussion, is that if $x \in \mathrm{MaxSpec}(R_E)$

then the corresponding residue field $E(x) = R_E/\mathfrak{m}_x$ is *finite* over E. It then makes sense to consider the *specialization* of ρ at x:

$$\rho_x : G_K \stackrel{\rho_E}{\to} \mathrm{GL}_n(R_E) \to \mathrm{GL}_n(E(x)).$$

Especially when E is finite over $\mathbf{Q}_p$, we visualize ρ as a "family" of p-adic representations $\{\rho_x\}$ with varying coefficient fields $E(x)$ of finite degree over E.

Remark 1.2. Note that each such ρ_x is continuous (and so *is* a p-adic representation of G_K) since x carries R into the valuation ring of $E(x)$ via a local map and ρ is continuous when R is given its local (i.e., max-adic) topology.

For a property $\mathbf{P}$ of (isomorphism classes of) G_K-representations over finite extensions of E and for any $x \in \mathrm{MaxSpec}(R_E)$, let $\mathbf{P}(x)$ denote the condition that ρ_x satisfies property $\mathbf{P}$. (In practice, $\mathbf{P}$ is always insensitive to finite scalar extension on the coefficient field over E.) It is useful to consider whether or not the locus

$$\mathbf{P}(R_E) = \{x \in \mathrm{MaxSpec}(R_E) \,|\, \mathbf{P}(x) \text{ holds}\}$$

is "analytic" in the sense that it is cut out by an ideal J of R_E. That is, for a Λ-algebra map $x : R \to E'$ to a finite extension E'/E, does ρ_x satisfy $\mathbf{P}$ if and only if $x(J) = 0$? A given ideal J in R_E satisfies this condition if and only if its radical does (since E' is reduced), so we may as well restrict attention to radical J. But since R_E is Jacobson, a radical ideal J in R_E is the intersection of the maximals over it, so in other words there is exactly one possibility for a radical J:

$$J_{\mathbf{P}} := \bigcap_{\mathbf{P}(x) \text{ holds}} \mathfrak{m}_x$$

where $\mathfrak{m}_x = \ker(x : R_E \to E(x))$. Note that if $\mathbf{P}(x)$ fails for all $x \in \mathrm{MaxSpec}(R_E)$ then $J_{\mathbf{P}} = (1)$ (either by logic, convention, or the utiliarian reason that it is consistent with what follows).

Turning this reasoning around, we take the above expression for $J_{\mathbf{P}}$ as a definition, so $V(J_{\mathbf{P}}) := \mathrm{Spec}(R_E/J_{\mathbf{P}})$ is the Zariski closure of the locus of $x \in \mathrm{MaxSpec}(R_E)$ such that $\mathbf{P}(x)$ holds. The analyticity question for $\mathbf{P}$ then amounts to the following question: *does every closed point of* $V(J_{\mathbf{P}})$ *satisfy* $\mathbf{P}$? It is by no means clear how one could answer this question, and in the early days of modularity lifting theorems this was a serious problem which had to be treated by ad hoc methods depending on the specific $\mathbf{P}$.

One of the big achievements of Kisin's introduction of integral p-adic Hodge theory into Galois deformation theory is to provide systematic techniques for proving an affirmative answer to this question for many interesting $\mathbf{P}$ involving conditions related to p-adic Hodge theory (e.g., crystalline with Hodge-Tate weights in the interval $[-2, 5]$). In any situation for which the $\mathbf{P}$-analyticity question has an affirmative answer, to exploit it one needs to answer a deeper question: *how can we analyze properties of* $R_E/J_{\mathbf{P}}$*, such as regularity, dimension, connectedness of spectrum, etc.*? Kisin's methods also gave a way to address this question. We will develop this for a special $\mathbf{P}$ that can be studied without the full force of p-adic Hodge theory: the "ordinary crystalline" deformation problem, to be defined later.

A key insight that underlies Kisin's strategy for answering both of these questions is to use $\mathbf{P}$ to define a new moduli problem on *arbitrary* R*-algebras* (forgetting the topological

structure of R) which is shown to be represented by a proper (even projective) R-scheme $\Theta : \mathscr{X}_{\mathbf{P}} \to \operatorname{Spec} R$ such that:

(1) the map $\Theta_E : X_{\mathbf{P},E} \to \operatorname{Spec}(R_E)$ obtained by inverting p (equivalently, localizing by $\Lambda \to E$) is a closed immersion whose image has as its closed points precisely the ones which satisfy $\mathbf{P}$ (so this closed subscheme, after killing nilpotents, recovers $J_{\mathbf{P}}$ and provides an affirmative answer to the $\mathbf{P}$-analyticity question),
(2) the Λ-scheme $X_{\mathbf{P}}$ is "formally smooth" in a sense we will make precise later. (In practice $X_{\mathbf{P}}$ is very far from being finite type over Λ, just like R itself, so we cannot naively carry over the notion of smooth morphism from algebraic geometry in terms of a Jacobian criterion.)

An important consequence of condition (2) is that the generic fiber $X_{\mathbf{P},E}$ is "formally smooth" over E, which is to say that it is regular and hence reduced. (In geometric language, this says that the rigid-analytic space over E arising from $R_E/J_{\mathbf{P}}$ in the open unit polydisc is smooth.) In particular, $X_{\mathbf{P},E} = V(R_E/J_{\mathbf{P}})$ (affine!), so the E-algebra $R_E/J_{\mathbf{P}}$ that we wish to understand is the coordinate ring of the affine generic fiber $X_{\mathbf{P},E}$ of the (typically non-affine!) moduli scheme $X_{\mathbf{P}}$ over R which we can try to study by moduli-theoretic reasoning. In fact, we will study the structure of the generic fiber over E by using moduli-theoretic considerations with the schemes $X_{\mathbf{P}} \bmod \pi R$ and $X_{\mathbf{P}} \bmod \mathfrak{m}_R$ which live in characteristic p!

Letting $I_{\mathbf{P}} = \ker(R \to R_E/J_{\mathbf{P}})$ be the ideal of the Zariski closure in $\operatorname{Spec} R$ of the $\mathbf{P}$-locus in $\operatorname{Spec} R_E$, the quotient $R/I_{\mathbf{P}}$ is reduced with generic $R_E/J_{\mathbf{P}}$. In practice we will think of $R/I_{\mathbf{P}}$ as an "integral parameter space for the property $\mathbf{P}$". In particular, the formal smoothness over Λ in (2) justifies viewing $X_{\mathbf{P}}$ as a "resolution of singularities" of $\operatorname{Spec}(R/I_{\mathbf{P}})$ (for which it has the same E-fiber).

2. Some commutative algebra and algebraic geometry

Before we launch into the definition and study of Kisin's moduli problems on R-algebras and their applications to the study of the $\mathbf{P}$-locus in $\operatorname{MaxSpec}(R_E)$, we digress to explain some general considerations in commutative algebra and algebraic geometry which will be used throughout his method. It will be clearer to carry out these general considerations now so that we will be ready for their applications later.

We consider the following general setup. Let (Λ, E, π, k, R) be as above, and let $f : X \to \operatorname{Spec} R$ be a proper R-scheme. There are two "reductions" of f that will be of interest: the reductions

$$\overline{f} : \overline{X} := X \bmod \pi \to \operatorname{Spec}(R/\pi R), \quad f_0 : X_0 := X \bmod \mathfrak{m}_R \to \operatorname{Spec} k$$

modulo πR and modulo $\mathfrak{m}_R$ respectively. In particular, X_0 is a proper (hence finite type) scheme over the residue field k. The quotient $R/\pi R$ is naturally a k-algebra (since $k = \Lambda/(\pi)$), but $\overline{X}$ is typically "huge" (not finite type) when thereby viewed as a k-scheme.

Since f is proper, it carries closed points of X into closed points of $\operatorname{Spec} R$. But there is only one closed point of the local $\operatorname{Spec} R$, and X_0 is closed in X, so we conclude that the closed points of X coincide with those of X_0. Moreover, if $x_0 \in X$ is a closed point then since it is closed in the scheme X_0 of finite type over k we see that the residue field $\kappa(x_0)$ of X at x_0 (or equivalently, of X_0 at x_0) is of *finite degree* over k,

Now assume that k is perfect (e.g., finite). Consider a closed point $x_0 \in X$, so $\kappa(x_0)/k$ is a finite separable extension. Let $\Lambda(x_0)$ be the unique (up to unique isomorphism) finite unramified extension of Λ with residue field $\kappa(x_0)/k$. The completed local ring $\mathscr{O}_{X,x_0}^{\wedge}$ is a Λ-algebra with residue field $\kappa(x_0)/k$, so by Hensel's Lemma it admits a unique structure of $\Lambda(x_0)$-algebra over its Λ-algebra structure. (This application of Hensel's Lemma crucially uses that we are working with the completed local ring and not the usual algebraic local ring $\mathscr{O}_{X,x_0}$; this latter Λ-algebra is typically not a $\Lambda(x_0)$-algebra in a compatible manner.)

Hypothesis $(*)$: assume that $\mathscr{O}_{X,x_0}^{\wedge} \simeq \Lambda(x_0)[\![T_1, \ldots, T_n]\!]$ as $\Lambda(x_0)$-algebras, for all closed points $x_0 \in X_0$.

This hypothesis can be checked by means of functorial criteria, and that is how it will be verified in later examples of interest. It follows from $(*)$ that the completion $\mathscr{O}_{X,x_0}^{\wedge}$ is regular, Λ-flat, and reduced modulo π for all x_0. These are the properties we will use to prove:

Proposition 2.1. *Under hypothesis $(*)$, the base change $\overline{X}$ over $R/(\pi)$ is reduced and the total space X is regular and Λ-flat.*

Proof. We first handle the Λ-flatness, and then turn to the claims concerning reducedness modulo π and regularity. The π-power torsion in $\mathscr{O}_X$ is a coherent ideal whose formation commutes with passage to stalks and completions thereof (by flatness of completion). For all closed points x_0, the completion of $\mathscr{O}_{X,x_0}$ is Λ-flat by inspection of its assumed structure. Hence, the π-power torsion ideal has vanishing stalks at all closed points, so it vanishes on an open subset of X which contains all closed points. Such an open subset must be the entire space, so X is Λ-flat.

By hypothesis $(*)$, each quotient $\mathscr{O}_{\overline{X},x_0}^{\wedge} = \mathscr{O}_{X,x_0}^{\wedge} \bmod \pi$ is reduced, so the proper scheme $\overline{X}$ over the complete local noetherian ring $\overline{R} = R/(\pi)$ has reduced local rings at the closed points. Thus, the coherent radical of the structure sheaf of $\mathscr{O}_{\overline{X}}$ has vanishing stalks at all closed points, so exactly as for the Λ-flatness above we conclude that $\overline{X}$ is reduced.

If the non-regular locus on X is closed then since the local rings on X at all closed points are regular (by inspection of their completions) it would follow that the non-regular locus is empty. That is, X is regular if the non-regular locus is closed. It remains to prove that the non-regular locus in X is Zariski-closed. The closedness of this locus in general locally noetherian schemes (and likewise for other properties defined by homological conditions) is a deep problem which was first systematically investigated by Grothendieck. His big discovery was that for a class of schemes called *excellent* the closedness always holds. He also proved that "most" noetherian rings which arise in practice are excellent.

We refer the reader to [4, Ch. 13] for an elegant development of the basic properties of excellence (including the definition!), and here we just record the main relevant points: excellence is a Zariski-local property, it is inherited through locally finite type maps, and every complete local noetherian ring (e.g., every field, as well as R above) is excellent. Hence, the scheme X is excellent, so its non-regular locus is Zariski closed. ■

Now we come to a very useful result which can be applied under the conclusions of the preceding proposition.

Lemma 2.2 (Reduced fiber trick)**.** *Let X be a Λ-flat R-scheme which is proper and for which $\overline{X} = X \bmod \pi$ is reduced. If X is connected and non-empty then $X_0 = X \bmod \mathfrak{m}_R$ and the generic fiber $X_E = X \otimes_\Lambda E$ are both connected and non-empty.*

In general (without connectedness hypotheses), there is a natural bijective correspondence between connected components C_0 of X_0 and C_E of X_E by the requirement that C_E is the E-fiber of the unique connected component of X with mod-$\mathfrak{m}_R$ fiber C_0.

Proof. Since X is non-empty and Λ-flat, X_E is non-empty. The theorem on formal functions, applied to the proper X over the complete local noetherian ring R, identifies the idempotents on X with those on X_0. In particular, X_0 is non-empty, and each connected component of X_0 uniquely lifts to a connected component of X. Hence, by passing to the connected components of X it suffices to prove that if X is connected then so is X_E.

The Λ-flatness of X implies that the ring $\mathscr{O}(X)$ of global functions on X injects into its localization $\mathscr{O}(X)[1/\pi] = \mathscr{O}(X)_E = \mathscr{O}(X_E)$ which is the ring of global functions on X_E. We assume that the latter contains an idempotent e and seek to prove $e = 0$ or $e = 1$. We can write $e = e'/\pi^n$ for a minimal $n \geq 0$ and a global function e' on X. If $n = 0$ then e' is idempotent on X and hence $e = e' \in \{0, 1\}$ since X is connected. Thus, we assume $n \geq 1$ and seek a contradiction to the minimality of n.

Since $e^2 = e$ on X_E, we can clear denominators (via Λ-flatness) to get ${e'}^2 = \pi^n e'$ on X. Thus, for $\overline{e}' = e' \bmod \pi$ we have $\overline{e}'^2 = 0$ on $\overline{X}$. But $\overline{X}$ is reduced, so $\overline{e}' = 0$ on $\overline{X}$. This says that e' is divisible by π locally on $\overline{X}$. Since X is Λ-flat, the local π-multiplier to get e' is unique and hence globalizes. That is, $e' = \pi e''$ for some $e'' \in \mathscr{O}(X)$. It follows that on X_E we have

$$e = \frac{e'}{\pi^n} = \frac{e''}{\pi^{n-1}},$$

contrary to the minimality of n. ∎

Inspired by the two preceding results, we are led to wonder: how can we ever verify Hypothesis $(*)$? We now present a functorial criterion.

Proposition 2.3. *Hypothesis $(*)$ holds if and only if for every artin local finite Λ-algebra B, $X(B) \to X(B/J)$ is surjective.*

Proof. Fix a closed point $x_0 \in X$ and let k'/k be a finite Galois extension which splits $k_0 := \kappa(x_0)$. Let $\Lambda_0 = \Lambda(x_0)$, and let Λ' be the finite unramified extension of Λ corresponding to k'/k. Thus,

$$\Lambda' \otimes_\Lambda \Lambda_0 \simeq \prod_{j : k_0 \to k'} \Lambda'_j \tag{2.1}$$

where Λ'_j denotes Λ' viewed as a $\Lambda(x_0)$-algebra via the unique Λ-embedding $\Lambda_0 \to \Lambda'$ lifting the k-embedding $j : k_0 \to k'$.

Recall from above that $\mathscr{O}^\wedge_{X,x_0}$ is canonically a Λ_0-algebra. Let B be an artin local finite Λ_0-algebra. A complete local noetherian Λ_0-algebra with residue field k_0 is a formal power series ring over Λ_0 if and only if it is Λ_0-flat with residue field k_0 and is regular modulo π. These properties hold if and only if the finite étale scalar extension by $\Lambda_0 \to \Lambda'$ yields the analogous properties using the residue field k', so it is equivalent to prove that this scalar

extension is a formal power series ring over Λ'. In particular, the functorial criterion for the latter condition is precisely that the natural map of sets

$$\mathrm{Hom}_{\Lambda_0}(\mathscr{O}^{\wedge}_{X,x_0}, B) \to \mathrm{Hom}_{\Lambda_0}(\mathscr{O}^{\wedge}_{X,x_0}, B/J) \tag{2.2}$$

is surjective for any artin local finite Λ'-algebra B with residue field of finite degree over k_0 and any square-zero ideal J in B.

We will reformulate this surjectivity in terms which are more easily related to the functor of points of X as we vary (B, J) with B an artin local finite Λ'-algebra. Since B is artin local and Λ-finite, the natural restriction map

$$\mathrm{Hom}_{\Lambda}(\mathscr{O}^{\wedge}_{X,x_0}, B) \to \mathrm{Hom}_{\Lambda}(\mathscr{O}_{X,x_0}, B) = X_{x_0}(B)$$

is bijective, where $X_{x_0}(B)$ denotes the set of Λ-maps $\mathrm{Spec}\, B \to X$ whose image is x_0. Using the Λ'-algebra structure on B and the canonical Λ'-algebra structure on $\mathscr{O}^{\wedge}_{X,x_0}$, we also have the alternative description

$$\mathrm{Hom}_{\Lambda}(\mathscr{O}^{\wedge}_{X,x_0}, B) = \mathrm{Hom}_{\Lambda'}(\Lambda' \otimes_{\Lambda} \mathscr{O}^{\wedge}_{X,x_0}, B) = \mathrm{Hom}_{\Lambda'}((\Lambda' \otimes_{\Lambda} \Lambda_0) \otimes_{\Lambda_0} \mathscr{O}^{\wedge}_{X,x_0}, B).$$

Using (2.1), this is identified with the disjoint union

$$\coprod_{j:k_0 \to k'} \mathrm{Hom}_{\Lambda'}(\Lambda' \otimes_{j,\Lambda_0} \mathscr{O}^{\wedge}_{X,x_0}, B) = \coprod_{j:k_0 \to k'} \mathrm{Hom}_{\Lambda_0}(\mathscr{O}^{\wedge}_{X,x_0}, B_j)$$

where B_j denotes B viewed as a Λ'-algebra via any $g \in \mathrm{Gal}(k'/k)$ lifting j on k_0.

The preceding identifications of Hom-sets are all functorial in B. In the final disjoint union above, as we vary through all pairs (B, J) with B an artin local finite Λ'-algebra and J a square-zero ideal in B, the simultaneous surjectivity of (2.2) for all pairs (B_j, J_j)'s is thereby identified with the surjectivity of the natural map

$$X_{x_0}(B) \to X_{x_0}(B/J)$$

as B varies through artin local finite Λ'-algebras. Recall that k'/k is an arbitrary but fixed finite Galois extension which splits $k_0 = \kappa(x_0)$. Thus, if $X(B) \to X(B/J)$ is surjective for all artin local finite Λ-algebras B and square-zero ideals $J \subset B$ then Hypothesis $(*)$ holds.

Conversely, if $(*)$ holds then for any such (B, J) we claim that $X(B) \to X(B/J)$ is surjective. Pick a point in $X(B/J)$. As a Λ-map $\mathrm{Spec}(B/J) \to X$ we claim that it hits a closed point x_0. Since B/J has residue field of finite degree over k, it suffices to show that this map lands in X_0 (as X_0 is a finite type k-scheme). Since the composite map $\mathrm{Spec}(B/J) \to \mathrm{Spec}(R)$ over Λ lifts a point of R valued in a *finite* extension of k, it suffices to check that the only such point is the evident one which kills $\mathfrak{m}_R$. Expressing $R/(\pi)$ as the quotient of a power series ring over Λ, it suffices to prove that the only local k-algebra map from $k[\![x_1, \ldots, x_m]\!]$ into a finite extension of k is "evaluation at the origin". This is verified by restriction to the k-subalgebras $k[\![x_i]\!]$ for $i = 1, \ldots, m$.

The chosen point in $X(B/J)$ identifies the residue field κ of B/J as a finite extension of the residue field k_0 at x_0, so B/J and hence B is thereby equipped with a natural structure of Λ_0-algebra over its Λ-algebra structure. The chosen point in $X(B/J)$ is thereby identified with a Λ-algebra map $\mathrm{Spec}(B/J) \to X$ hitting x_0 and lifting the specified extension structure

κ/k_0. This map corresponds to a local Λ-algebra map $\mathscr{O}_{X,x_0} \to B/J$ lifting $k_0 \to \kappa$, which in turn uniquely factors through such a local Λ-map

$$\mathscr{O}^{\wedge}_{X,x_0} \to B/J.$$

This latter map is a Λ_0-algebra map (as may be checked on residue fields). By $(*)$, the completion $\mathscr{O}^{\wedge}_{X,x_0}$ is a formal power series ring over Λ_0, so its map to B/J lifts to a Λ_0-algebra map $\mathscr{O}^{\wedge}_{X,x_0} \to B$. Running the procedure in reverse, this gives a Λ-map $\operatorname{Spec} B \to X$ which lifts the chosen point in $X(B/J)$. ■

Remark 2.4. The criterion in Proposition 2.3 is what was used in the verification of power series properties in the preceding lecture on the case $\ell \neq p$. In Kisin's papers, he expresses things in terms of a much more general theory of formal smoothness for maps of topological rings, and he thereby invokes some very deep results of Grothendieck in this theory.

For example, a noetherian algebra over a field of characteristic 0 is formally smooth (for the discrete topology) over that field if and only if it is regular. We will speak in the language of regularity and avoid any need for the theory of formal smoothness because we will appeal to general results in the theory of excellence (as was done in the proof of Proposition 2.1).

The reader who is interested in reading up on the general theory of formal smoothness (such as its flatness aspects) should look at [5, Ch. 28], §17.5 in EGA IV_4, and §19–22 (esp. 19.7.1 and 22.1.4) in Chapter 0_{IV} of EGA. Certainly if one goes deeper into Kisin's techniques (beyond the "ordinary crystalline" deformation condition to be considered below) then it becomes important to use formal smoothness techniques in the generality considered by Kisin.

The final topic we take up in this section is the algebro-geometric problem of giving a convenient criterion to prove that the proper map

$$f_E : X_E \to \operatorname{Spec}(R_E)$$

is a closed immersion. More specifically, we want to give a criterion involving points valued in *finite* E-algebras C. Keep in mind that even though R_E is a gigantic E-algebra in general, it is Jacobson and its maximal ideals have residue field of finite E-degree. In particular, the artinian quotients of R_E at its maximal ideals are examples of such E-algebras C. The same goes for the R_E-proper X_E at its closed points (which lie over $\operatorname{MaxSpec}(R_E)$, due to the properness of f_E).

Proposition 2.5. *If f_E is injective on C-valued points for all E-finite C then f_E is a closed immersion.*

Proof. We will first prove that f_E is a finite map (i.e., X_E is the spectrum of a finite R_E-algebra), so then we can use Nakayama's Lemma to check the closed immersion property. Since f_E is proper, it suffices to prove that it is quasi-finite. For any map of finite type between noetherian schemes, the locus of points on the source which are isolated in their fibers (i.e., the "quasi-finite locus") is an open set: this is a special case of semi-continuity of fiber dimension. Thus, if f_E has finite fibers over $\operatorname{MaxSpec}(R_E)$ then the open quasi-finite locus of f_E contains all closed points of X_E (as these are precisely the points over $\operatorname{MaxSpec}(R_E)$, due to properness of f_E). But X_E is a Jacobson scheme since it is finite type

over the Jacobson ring R_E, and (as for any noetherian topological space) the only open set in X_E which contains all closed points is the entire space. Hence, if f_E has finite fibers over $\mathrm{MaxSpec}(R_E)$ then f_E is a quasi-finite and therefore *finite* map.

Letting C vary through the finite extension fields E'/E, the injectivity of f_E on E'-valued points implies that the fiber of f_E over each $y \in \mathrm{MaxSpec}(R_E)$ has only finitely many closed points. (Here we use that $f_E^{-1}(y)$ is closed in X_E and is of finite type over $E(y)$, with $E(y)$ finite over E.) But a scheme of finite type over a field has finitely many closed points if and only if it is finite. Thus, f_E indeed has finite fibers over $\mathrm{MaxSpec}(R_E)$. This argument even shows that such fibers have *at most one* physical point (since if a fiber contains two distinct points $x' \neq x$ then using E' containing $E(x)$ and $E(x')$ makes f_E fail to be injective on E'-valued points).

Now consider the finite map $f_E : X_E \to \mathrm{Spec}\, R_E$. To prove that the corresponding module-finite map of coordinate rings is surjective (so f_E is a closed immersion), it suffices to check surjectivity after localizing at maximal ideals of R_E. By Nakayama's Lemma, it is equivalent to check that the scheme-theoretic fiber $\mathrm{Spec}\, C \to \mathrm{Spec}\, E(y)$ of f_E over each $y \in \mathrm{MaxSpec}(R_E)$ satisfies $C = 0$ or $C = E(y)$. The two composite maps

$$\mathrm{Spec}(C \otimes_{E(y)} C) \rightrightarrows \mathrm{Spec}\, C \to \mathrm{Spec}\, E(y)$$

coincide, so for the E-finite algebra $C' = C \otimes_{E(y)} C$ we see that the composites

$$\mathrm{Spec}(C') \rightrightarrows \mathrm{Spec}\, C = X_E \times_{\mathrm{Spec}(R_E)} \mathrm{Spec}(E(y)) \hookrightarrow X_E$$

have the same composition with $f_E : X_E \to \mathrm{Spec}(R_E)$. By hypothesis, f_E is injective on C'-valued points! Hence, the projections $\mathrm{Spec}(C') \rightrightarrows \mathrm{Spec}(C)$ coincide, which is to say that the two inclusions $C \rightrightarrows C' = C \otimes_{E(y)} C$ coincide. This easily forces $C = E(y)$ if $C \neq 0$ (by consideration of an $E(y)$-basis of C containing 1). ■

3. The ordinary crystalline deformation problem

Now assume that k is finite! Let E'/E be a finite extension and Λ' its valuation ring. Fix a continuous representation $\rho : G_K \to \mathrm{GL}_2(E')$ with cyclotomic determinant χ. We already know what it means to say that ρ is *ordinary*: this means that there is a G_K-equivariant quotient line with action by an unramified character η. Such a quotient line is unique, as the $\Lambda^\times$-valued $\det \rho = \chi$ is ramified, so it is equivalent to say that ρ admits an I_K-equivariant quotient line with trivial I_K-action. This notion of ordinarity can be expressed in terms of a $\mathrm{GL}_2(\Lambda')$-valued conjugate of ρ by using saturated Λ'-lines.

In terms of a G_K-stable Λ'-lattice, we get an upper-triangular form for ρ, or equivalently for $\rho|_{I_K}$, and this extension structure identifies $\rho|_{I_K}$ with a class in

$$\mathrm{H}^1(I_K, \Lambda'(1)) = \varprojlim \mathrm{H}^1(I_K, (\Lambda'/(p^n))(1))$$

with $\Lambda'/(p^n)$ a finite free $\mathbf{Z}/(p^n)$-module since $[k : \mathbf{F}_p]$ is now assumed to be finite. This rank is equal to the $\mathbf{Z}_p$-rank of the finite free $\mathbf{Z}_p$-module Λ'. By computing with a $\mathbf{Z}_p$-basis of Λ', the natural map

$$\Lambda' \otimes_{\mathbf{Z}_p} \mathrm{H}^1(I_K, \mu_{p^n}) \to \mathrm{H}^1(I_K, (\Lambda'/(p^n))(1))$$

is an isomorphism, and we can pass this tensor product through an inverse limit to get

$$\mathrm{H}^1(I_K, \Lambda'(1)) = \Lambda' \otimes_{\mathbf{Z}_p} \mathrm{H}^1(I_K, \mathbf{Z}_p(1)).$$

The "crystalline" property of ρ is now going to be defined in terms of a description of $\mathrm{H}^1(I_K, \mathbf{Z}_p(1))$ and the identification of $\rho|_{I_K}$ as a class in $\Lambda' \otimes_{\mathbf{Z}_p} \mathrm{H}^1(I_K, \mathbf{Z}_p(1))$. (After making the definition, we will make it more concrete in terms of matrices.) We note at the outset that the definition we will give is in fact equivalent to a special case of a general notion of "crystalline" defined in p-adic Hodge theory, but we have avoided any discussions of p-adic Hodge theory and so will likewise have no need to delve further into the justification for our choice of terminology. A reader who pursues the subject in greater depth will eventually meet the general concept of "crystalline", but it is logically unnecessary for our purposes.

Let K' denote the completion of the maximal unramified extension K^{un}, so

$$K' = W(\overline{\mathbf{F}}) \otimes_{W(\mathbf{F})} K$$

where $\mathbf{F}$ is the finite residue field of K. Thus, $I_K = G_{K'}$ and $\mathscr{O}_{K'}$ is a complete discrete valuation ring with uniformizer given by one for $\mathscr{O}_K$. In particular, $1 + \mathfrak{m}_{K'}$ is p-adically separated and complete as a multiplicative $\mathbf{Z}_p$-module. By Kummer theory,

$$\mathrm{H}^1(G_{K'}, \mathbf{Z}_p(1)) = \varprojlim K'^{\times}/K^{\mathrm{un}\times p^n} = \mathbf{Z}_p \times (1 + \mathfrak{m}_{K'}) \tag{3.1}$$

where the $\mathbf{Z}_p$-factor corresponds to powers of a fixed uniformizer of K (or K'). This direct product decomposition is not canonical: the direct factor of $\mathbf{Z}_p$ depends on a choice of uniformizer. However, the "$\mathbf{Z}_p$-hyperplane" of 1-units $1 + \mathfrak{m}_{K'}$ (a multiplicative $\mathbf{Z}_p$-module) is canonical.

Definition 3.1. The ordinary representation $\rho : G_K \to \mathrm{GL}_2(\Lambda')$ is *crystalline* if its class in $\mathrm{H}^1(I_K, \Lambda'(1))$ lies in the Λ'-hyperplane

$$(1 + \mathfrak{m}_{K'}) \otimes_{\mathbf{Z}_p} \Lambda'.$$

Equivalently, $\rho : G_K \to \mathrm{GL}_2(E')$ corresponds to a class in $\mathrm{H}^1(I_K, E'(1))$ lying in the E'-hyperplane $(1 + \mathfrak{m}_{K'}) \otimes_{\mathbf{Z}_p} E'$.

In view of the formulation over E', the crystalline condition is intrinsic to the E'-linear representation space for G_K, so it does not depend on a specific choice of G_K-stable Λ'-lattice.

Example 3.2. The concrete meaning of the crystalline condition is as follows. In terms of a choice of G_K-stable Λ'-lattice, consider $\rho|_{I_K} \bmod p^n$ for each $n \geq 1$. This is upper triangular unipotent, with upper-right entry given by a Λ'-linear combination of 1-cocycles $g \mapsto g(u^{1/p^n})/u^{1/p^n}$ on I_K, with $u \in \mathscr{O}_{K'}^{\times}$. Loosely speaking, Kummer theory shows that $\rho|_{I_K}$ is given by a Tate-curve type of construction with Λ'-coefficients, and the crystalline condition is that the "q-parameter" can be chosen to be a unit (or equivalent a 1-unit, since $\overline{\mathbf{F}}^{\times}$ is uniquely p-divisible).

Now we fix a residual representation $\rho_0 : G_K \to \mathrm{GL}_2(k)$ with cyclotomic determinant, and we seek to study its ordinary crystalline deformations with coefficients in p-adic fields or valuation rings thereof. In order to make good sense of these notions in deformation theory, we need to generalize the definition of "ordinary crystalline" to the case of more general coefficients.

Remark 3.3. We will later need to consider coefficients in rings like $R[T]$ that are *not* $\mathfrak{m}_R$-adically separated and complete, so this will create some delicate problems when we work

with finite type R-schemes such as $\mathbf{P}^1_R$ that we want to use in the construction of a moduli scheme for an "ordinary crystalline" deformation problem. The creative use of proper (generally non-finite) R-schemes in Galois deformation theory is one of the innovations introduced by Kisin's work.

Consider the universal deformation $\rho : G_K \to \mathrm{GL}_2(R)$ or universal framed deformation $\rho^{\square} : G_K \to \mathrm{GL}_2(R^{\square})$ of ρ_0 with cyclotomic determinant. We want to study the locus of "ordinary crystalline points" in $\mathrm{MaxSpec}(R_E)$. The formalism for this study will not really use the universality at all, so to keep the picture clear we now consider *any* continuous $\rho : G_K \to \mathrm{GL}_2(R)$ as at the outset such that $\det\rho$ is cyclotomic, and we continue to assume that the residue field k of R is *finite*. For any R-algebra A, let

$$\rho_A : G_K \to \mathrm{GL}_2(A)$$

denote the composition of ρ with $R \to A$ on matrix entries. Note that there is no meaningful continuity condition for ρ_A for general A, since we are not assuming that A carries an interesting topology compatible with the one on R.

In the special case that A is an $R/\mathfrak{m}_R^n$-algebra, ρ_A is continuous for the discrete topology on A (and the Krull topology on G_K) since $\rho \bmod \mathfrak{m}_R^n$ is continuous for the discrete topology on $R/\mathfrak{m}_R^n$. Beware that if we take $A = R/(p^n)$ with the discrete topology then ρ_A is typically *not* continuous. It will therefore be important that we can work modulo powers of the maximal ideal of R and bootstrap back up to geometric objects over R via limit procedures.

Example 3.4. Our work with ρ_A for $R/\mathfrak{m}_R^n$-algebras A will involve some Galois cohomology with A-coefficients viewed discretely, so we record here the useful fact that for any $\mathbf{Z}/(p^n)$-module M viewed discretely (such as $M = A$) the natural map

$$M \otimes_{\mathbf{Z}_p} \mathrm{H}^1(I_K, \mathbf{Z}_p(1)) \to \mathrm{H}^1(I_K, M(1))$$

is an isomorphism.

To prove this, we can use direct limits in M to reduce to the case when M is a finitely generated $\mathbf{Z}_p$-module. Hence, M is a finite direct sum of modules of the form $\mathbf{Z}/(p^r)$ with $r \leq n$, so it suffices to treat the case $M = \mathbf{Z}/(p^r)$. Then the assertion is that the natural map

$$\mathrm{H}^1(I_K, \mathbf{Z}_p(1))/(p^r) \to \mathrm{H}^1(I_K, \mu_{p^r})$$

is an isomorphism for all $r \geq 1$. For K' denoting the discretely-valued completion of K^{un} we have $I_K = G_{K'}$ and $1 + \mathfrak{m}_{K'}$ is p-adically separated and complete (as a multiplicative $\mathbf{Z}_p$-module), so Kummer theory and the description of $\mathrm{H}^1(I_K, \mathbf{Z}_p(1))$ in (3.1) yields the result.

In the special case that A is the valuation ring of a finite extension of E, we have defined what it means to say that ρ_A is ordinary crystalline (in Definition 3.1). That definition involved the p-adic topology of the valuation ring. We wish to define this concept for $R/\mathfrak{m}_R^n$-algebras A, avoiding any use of nontrivial topologies on rings.

Definition 3.5. Let A be an R-algebra killed by $\mathfrak{m}_R^n$ for some $n \geq 1$. The representation ρ_A of G_K on $V_A := A^2$ is *ordinary crystalline* if there is a G_K-stable A-submodule $L_A \subset V_A$ such that L_A and V_A/L_A are locally free of rank 1 (equivalently, projective of rank 1) and:

(1) the I_K-action on V_A/L_A is trivial, or equivalently the continuous action of G_K on V_A/L_A is through an unramified character $\eta : G_K \to A^\times$ and on L_A is through $\chi\eta^{-1}$;
(2) under the "valuation" map $K^{\mathrm{un}\times} \twoheadrightarrow \mathbf{Z}$, the class in

$$\mathrm{H}^1(I_K, A(1)) = \mathrm{H}^1(I_K, \mathbf{Z}_p(1)) \otimes_{\mathbf{Z}_p} A = (K^{\mathrm{un}\times}/(K^{\mathrm{un}\times})^{p^n}) \otimes_{\mathbf{Z}/(p^n)} A$$

describing $\rho_A|_{I_K}$ is carried to $0 \in A$ (i.e., the class arises from integral units of K^{un}).

We call the A-line $L_A \subset V_A$ an *ordinary crystalline structure* on ρ_A.

Remark 3.6. In the "crystalline" condition (2) in this definition, we have invoked the cohomology computation in Example 3.4. Also, in general there may be more than one choice of L_A (if any exist at all!). For example, if $A = k$ and ρ_k has trivial G_K-action (in particular, the mod-p cyclotomic character is trivial) then *every* line in $V_k = k^2$ is an ordinary crystalline structure on ρ_k.

A fundamental insight of Kisin is that rather than trying to parameterize deformations which admit an additional structure (such as an ordinary crystalline structure) that may not be unique, it is better to parameterize the space of *pairs* consisting of a deformation *equipped with* such an additional structure. To make reasonable sense of a parameter space for such enhanced objects, we will have to leave the framework of complete local noetherian rings and instead work with certain proper schemes over such rings.

The property in (2) in the preceding definition makes sense as a condition on classes in $\mathrm{H}^1(I_K, M(\chi))$ for any discrete $\mathbf{Z}/(p^n)$-module M for any $n \geq 1$. (Note that $M(\chi)$ is a discrete I_K-module, and even a discrete G_K-module.)

Definition 3.7. For any discrete $\mathbf{Z}/(p^n)$-module M equipped with a continuous unramified G_K-action (not necessarily trivial), the subgroup

$$\mathrm{H}^1_{\mathrm{crys}}(K, M(\chi)) \subset \mathrm{H}^1(K, M(\chi))$$

consists of classes whose restriction to $\mathrm{H}^1(I_K, M(\chi)) = M \otimes_{\mathbf{Z}_p} \mathrm{H}^1(I_K, \mathbf{Z}_p(1))$ is killed by the "valuation" mapping $\mathrm{H}^1(I_K, \mathbf{Z}_p(1)) \to \mathbf{Z}_p$ defined by Kummer theory. (In other words, the I_K-restriction is an M-linear combination of classes in $\mathrm{H}^1(I_K, \mathbf{Z}_p(1))$ arising from integral units of the completion of K^{un}.)

It is immediate from the definition that if $M = \varinjlim M_i$ for a directed system of unramified discrete $(\mathbf{Z}/(p^n))[G_K]$-modules M_i then the equality

$$\varinjlim \mathrm{H}^1(K, M_i(\chi)) = \mathrm{H}^1(K, M(\chi))$$

carries $\varinjlim \mathrm{H}^1_{\mathrm{crys}}(K, M_i(\chi))$ isomorphically onto $\mathrm{H}^1_{\mathrm{crys}}(K, M(\chi))$. In other words, the formation of $\mathrm{H}^1_{\mathrm{crys}}(K, M(\chi))$ is compatible with direct limits in M. This will be very useful for reducing some general assertions to the special case of M which are $\mathbf{Z}_p$-finite (whereas in applications we will need to work with M which are not $\mathbf{Z}_p$-finite, such as $(R/\mathfrak{m}_R^n)[t]$).

Example 3.8. Isomorphism classes of pairs (ρ_A, L_A) as in Definition 3.5 correspond to elements in $\mathrm{H}^1_{\mathrm{crys}}(K, A(1))$.

Example 3.9. It is important to link Definition 3.5 and Definition 3.1. For $\rho : G_K \to \mathrm{GL}_2(\Lambda')$ as in Definition 3.1 we claim that it is ordinary crystalline in that initial sense (which can be checked over the fraction field E' of Λ') if and only if the artinian quotients $\rho \bmod \pi^n\Lambda'$ of ρ are ordinary crystalline in the sense of Definition 3.5 with $R = \Lambda'$ (i.e., each $\rho \bmod \pi^n\Lambda'$ admits an ordinary crystalline structure).

It is obvious that if ρ is ordinary crystalline in the initial sense then each artinian quotient $\rho \bmod \pi^n\Lambda'$ admits an ordinary crystalline structure. To go in reverse, suppose that every such artinian quotient admits an ordinary crystalline structure. Such structures are not unique in general, but since k is *finite* there are only finitely many such structures for each $n \geq 1$. These finite non-empty sets form an inverse system in an evident manner, and so the inverse limit is non-empty. (This is an elementary fact since the inverse system is indexed by positive integers and not a general index set.)

An element of the inverse limit is precisely the data of a saturated G_K-stable Λ'-line L in ρ such that (i) $\rho \bmod L$ has trivial I_K-action (as may be checked modulo π^n for all $n \geq 1$), and (ii) the class in $\mathrm{H}^1(I_K, \Lambda'(1)) = \Lambda' \otimes_{\mathbf{Z}_p} \mathrm{H}^1(I_K, \mathbf{Z}_p(1))$ corresponding to $(\rho|_{I_K}, L)$ has image under the "valuation mapping" $\mathrm{H}^1(I_K, \Lambda'(1)) \to \Lambda'$ which vanishes (as this can also be checked modulo π^n for all $n \geq 1$). The conditions (i) and (ii) say exactly that ρ is ordinary crystalline in the sense of Definition 3.1.

The proof of a later "formal smoothness" result over Λ will rest on:

Lemma 3.10. *For any $n \geq 1$, the functor $M \rightsquigarrow \mathrm{H}^1_{\mathrm{crys}}(K, M(\chi))$ on discrete unramified $(\mathbf{Z}/(p^n))[G_K]$-modules is right-exact.*

This is analogous to the fact that $\mathrm{H}^1(I_K, M(\chi)) = M \otimes_{\mathbf{Z}_p} \mathrm{H}^1(I_K, \mathbf{Z}_p(1))$ (see Example 3.4) is right-exact in discrete p^n-torsion abelian groups M with trivial I_K-action.

Proof. By discreteness we can express any M as a direct limit of $\mathbf{Z}_p$-finite G_K-submodules, so any right exact sequence in M's is obtained as a direct limit of right-exact sequences of $\mathbf{Z}_p$-finite object. Thus, the compatibility with direct limits in M reduces the problem to right-exactness for M which are finite abelian p-groups.

There are two ways to settle the finite case. In [3, 2.4.2], Kisin does some work with cocycles to derive an explicit description of $\mathrm{H}^1_{\mathrm{crys}}(K, M(\chi))$ which makes the right-exactness evident by inspection. This is definitely the most elementary way to proceed.

For the reader who doesn't like cocycle arguments and is familiar with the fppf topology, here is an alternative explanation in such terms. This explanation is longer, but may be seen as more conceptual (and clarifies the role of finiteness of the residue field).

The finite discrete G_K-module $M(\chi)$ has unramified Cartier dual, so it is the generic fiber of a unique finite flat $\mathscr{O}_K$-group scheme $M(\chi)'$ with étale Cartier dual, and $M(\chi)'$ is functorial in M. If

$$0 \to M_1 \to M_2 \to M_3 \to 0$$

is an exact sequence of such unramified G_K-modules then the complex of finite flat $\mathscr{O}_K$-group schemes

$$0 \to M_1(\chi)' \to M_2(\chi)' \to M_3(\chi)' \to 0$$

is short exact (in particular, short exact as abelian sheaves for the fppf topology over $\mathscr{O}_K$), as may be checked using the finite étale Cartier duals.

It can be shown by a Kummer-theoretic argument in the fppf topology that

$$\mathrm{H}^1_{\mathrm{crys}}(K, M(\chi)) = \mathrm{H}^1_{\mathrm{fppf}}(\mathscr{O}_K, M(\chi)'),$$

so it suffices to show that $\mathrm{H}^1_{\mathrm{fppf}}(\mathscr{O}_K, M(\chi)')$ is right exact in M. Equivalently, a short exact sequence in M induces a right-exact sequence in $\mathrm{H}^1_{\mathrm{fppf}}(\mathscr{O}_K, M(\chi)')$. The long-exactness of fppf cohomology then does the job provided that $\mathrm{H}^2_{\mathrm{fppf}}(\mathscr{O}_K, M(\chi)') = 0$ for any finite abelian p-group M equipped with an unramified continuous G_K-action.

The filtration by $\{p^m M\}$ reduces us to the case when M is p-torsion. If $r = \dim_{\mathbf{F}_p} M$ and k'/k is a finite Galois extension which splits M then for the corresponding finite unramified extension K'/K we have that $M(\chi)'_{\mathscr{O}_{K'}} = \mu_p^r$. Thus, we have an $\mathscr{O}_K$-subgroup inclusion

$$M(\chi)' \hookrightarrow \mathrm{Res}_{\mathscr{O}_{K'}/\mathscr{O}_K}(M(\chi)'_{\mathscr{O}_{K'}}) = \mathrm{Res}_{\mathscr{O}_{K'}/\mathscr{O}_K}(\mu_p^r),$$

where $\mathrm{Res}_{\mathscr{O}_{K'}/\mathscr{O}_K}$ denotes Weil restriction of scalars. This latter operation represented push-forward at the level of fppf sheaves, and it is an exact functor because $\mathscr{O}_K \to \mathscr{O}_{K'}$ is finite étale (and hence a split covering étale-locally over $\mathrm{Spec}(\mathscr{O}_K)$).

We conclude that $M(\chi)'$ is contained in the $\mathscr{O}_K$-group scheme $T' = \mathrm{Res}_{\mathscr{O}_{K'}/\mathscr{O}_K}(\mathbf{G}_m^r)$ which is an $\mathscr{O}_K$-torus (as we see by working étale-locally to split the covering $\mathrm{Spec}(\mathscr{O}_{K'}) \to \mathrm{Spec}(\mathscr{O}_K)$). Hence, we have a short exact sequence

$$1 \to M(\chi)' \to T' \to T'' \to 1$$

where $T'' := T'/M(\chi)'$ is another $\mathscr{O}_K$-torus. Thus, using the resulting long exact sequence in fppf cohomology, to prove $\mathrm{H}^2_{\mathrm{fppf}}(\mathscr{O}_K, M(\chi)') = 0$ it suffices to prove that $\mathrm{H}^2_{\mathrm{fppf}}(\mathscr{O}_K, T')$ vanishes and that any $\mathscr{O}_K$-torus (such as T'') has vanishing degree-1 cohomology. For the latter, first recall that degree-1 cohomology with affine coefficients classifies fppf torsors, so the degree-1 vanishing amounts to the triviality of such torsors when the coefficients are smooth and affine with connected fibers (such as a torus). To build a section splitting such a torsor over $\mathscr{O}_K$ it suffices (by smoothness of the coefficients, and the henselian property for $\mathscr{O}_K$) to find a section over the residue field. That is, we are reduced to proving the vanishing of degree-1 cohomology over k with coefficients in a smooth connected affine group (such as a torus). This is Lang's theorem, since k is finite.

Finally, to prove that $\mathrm{H}^2_{\mathrm{fppf}}(\mathscr{O}_K, T')$ is trivial, by using the definition of T' this amounts to vanishing of

$$\mathrm{H}^2_{\mathrm{fppf}}(\mathscr{O}_K, \mathrm{Res}_{\mathscr{O}_{K'}/\mathscr{O}_K}(\mathbf{G}_m)).$$

The exactness of the Weil restriction functor in this case implies (by a δ-functor argument) that

$$\mathrm{H}^j_{\mathrm{fppf}}(\mathscr{O}_K, \mathrm{Res}_{\mathscr{O}_{K'}/\mathscr{O}_K}(\cdot)) \simeq \mathrm{H}^j_{\mathrm{fppf}}(\mathscr{O}_{K'}, \cdot)$$

for all $j \geq 0$. Hence, we just have to check the vanishing of $\mathrm{H}^2_{\mathrm{fppf}}(\mathscr{O}_{K'}, \mathbf{G}_m)$ By Grothedieck's work on Brauer groups, this is identified with $\mathrm{Br}(\mathscr{O}_{K'})$, and since $\mathscr{O}_{K'}$ is henselian this in turn is identified with $\mathrm{Br}(k')$. But k' is *finite*, so its Brauer group vanishes. ■

Now we can prove the main existence result for a proper (even projective) R-scheme that "classifies" ordinary-crystalline pushforwards of ρ.

Theorem 3.11. *For each $n \geq 1$, the functor on $R/\mathfrak{m}_R^n$-algebras given by*

$$A \rightsquigarrow \{L_A \subset V_A \,|\, L_A \text{ is an ordinary crystalline structure on } \rho_A\}$$

is represented by a closed subscheme $X_n \subseteq \mathbf{P}^1_{R/\mathfrak{m}_R^n}$.

There is a unique closed subscheme $X \subseteq \mathbf{P}^1_R$ such that $X \bmod \mathfrak{m}_R^n = X_n$ for all $n \geq 1$.

Proof. By the universal property of the projective line, $P_n := \mathbf{P}^1_{R/\mathfrak{m}_R^n}$ represents the functor carrying any $R/\mathfrak{m}_R^n$-algebra A to the set of locally free A-submodules $L_A \subset V_A = A^2$ of rank 1 such that V_A/L_A is also locally free of rank 1. Over P_n, consider the G_K-action on $\mathscr{O}_{P_n}^2$ defined by $\rho \bmod \mathfrak{m}_R^n$ and the $R/\mathfrak{m}_R^n$-algebra structure on $\mathscr{O}_{P_n}$.

For each $g \in G_K$ and any A-point of P_n, the condition that A-pullback of the g-action on $\mathscr{O}_{P_n}^2$ preserves the corresponding A-line in A^2 is represented by a closed subscheme Z_g of P_n. To prove this, we may work Zariski-locally on P_n so that the universal line subbundle is free and extends to a basis of $\mathscr{O}^2$. Then the vanishing of the resulting "lower left matrix entry" function over the open in the base is what cuts out the g-stability condition over such an open locus in P_n. These closed loci agree on overlaps and glue to the desired closed subscheme Z_g of P_n attached to g. Thus, the closed subscheme $Z_n = \cap_g Z_g \subseteq P_n$ representing the condition of G_K preserving the universal line subbundle of $\mathscr{O}^2$.

Consider the character $\eta_n : G_K \to \mathscr{O}(Z_n)^\times$ describing the G_K-action on the universal line subbundle over Z_n. The Zariski-closed conditions $\eta_n(g) = 1$ for all $g \in I_K$ cut out a closed subscheme $Z_n' \subseteq Z_n$ which represents the additional condition that the universal line subbundle is not only G_K-stable but has unramified G_K-action. In other words, Z_n' represents the functor of "ordinary structures" on ρ_n.

Over Z_n', consider the further condition that the ordinary structure is crystalline. That is, for an A-point of Z_n', we consider the property that the resulting A-line L_A in ρ_A is an ordinary crystalline structure. The map

$$\mathrm{H}^1(I_K, A(1)) \to A$$

defined by the valuation $K^{\mathrm{un}\times} \to \mathbf{Z}$ carries the class of (ρ_A, L_A) to an element $a \in A$, and this construction is functorial in A. Hence, by (the proof of) Yoneda's Lemma it defines a global function h_n on Z_n'. The zero scheme of h_n on Z_n' is clearly the desired X_n.

Having constructed $X_n \subset P_n$ for each $n \geq 1$, the behavior of moduli schemes with respect to base change implies that the isomorphism

$$P_n \simeq P_{n+1} \otimes_{R/\mathfrak{m}_R^{n+1}} (R/\mathfrak{m}_R^n)$$

carries X_n over to $X_{n+1} \bmod \mathfrak{m}_R^n$. In other words, $\{X_n\}$ is a system of compatible closed subschemes of the system $\{P_n\}$ of infinitesimal fibers of the proper morphism $\mathbf{P}^1_R \to \operatorname{Spec} R$ over the complete local noetherian ring R. Now comes the deepest step: by Grothendieck's "formal GAGA" (EGA III_1, §5), if R is any complete local noetherian ring and P is any proper R-scheme, then the functor

$$Z \rightsquigarrow \{Z \bmod \mathfrak{m}_R^n\}$$

from closed subschemes of P to systems of compatible closed subschemes of the infinitesimal fibers of P over $\operatorname{Spec} R$ is a *bijection*. (Even for $P = \mathbf{P}^1_R$ this is not obvious, and it fails

miserably if we consider the affine line instead of the projective line.) Thus, we get the existence and uniqueness of the desired X. ■

Remark 3.12. Although each infinitesimal fiber X_n of X over $\operatorname{Spec} R$ has moduli-theoretic meaning for points valued in arbitrary $R/\mathfrak{m}_R^n$-algebras, we do *not* claim that X has a convenient moduli-theoretic meaning for its points valued in arbitrary R-algebras. In particular, $X_E = X \otimes_R R_E$ has no easy interpretation.

Nonetheless, it is X_E which will be of most interest to us. Thus, to work with X_E we need a way to understand its properties by studying the X_n's. This problem will be taken up in the next section.

4. Properties and applications of the ordinary crystalline moduli scheme

The construction of the proper R-scheme $f : X \to \operatorname{Spec} R$ is indirect, as formal GAGA is very abstract, but we can artfully use the construction to infer global properties of X which will be especially useful for the study of X_E. Our analysis rests on the following hypothesis which is in force throughout this section (unless we say otherwise):

Assume that $\det \rho : G_K \to R^\times$ *is cyclotomic and that* ρ *is the universal framed cyclotomic-determinant deformation ring of its reduction* ρ_k*. If* ρ_k *has only scalar endomorphisms, we also allow that* (ρ, R) *is the universal deformation of* ρ_k *with cyclotomic determinant.*

Proposition 4.1. *The* Λ*-scheme* X *is regular and flat, and* $X \bmod \pi$ *is reduced.*

Proof. By Proposition 2.1, Lemma 2.2, and Proposition 2.3, it suffices to prove that $X(B) \to X(B/J)$ is surjective for every artin local finite Λ-algebra B. Choose such a B, and let k' be its residue field, so B is canonically an algebra over $\Lambda' = W(k') \otimes_{W(k)} \Lambda$. It is harmless to make the finite étale scalar extension by $\Lambda \to \Lambda'$ throughout (this is compatible with the formation of X) to reduce to the case $k' = k$.

The maximal ideal of B is nilpotent, say with vanishing nth power for some $n \geq 1$, so the map of interest on points of X coincides with the analogue for X_n. Thus, the task is to show that if $\overline{L}$ is an ordinary crystalline structure on $\rho_{B/J}$ then it lifts to one on ρ_B. Let $x_0 \in X$ be the closed k-point corresponding to the specialization $(\rho_k, \overline{L}_k)$ of $(\rho_{B/J}, \overline{L})$ over the residue field k of B. Our problem is equivalent to showing that any local Λ-algebra map $\mathscr{O}^\wedge_{X,x_0} \to B/J$ lifts to B. Thus, it is sufficient (and even necessary) to prove that $\mathscr{O}^\wedge_{X,x_0}$ is a formal power series ring over Λ. To do this, we need to give a *deformation-theoretic* interpretation of this completion.

Since x_0 is a closed point, $R \to \mathscr{O}^\wedge_{X,x_0}$ is a local map and its reduction modulo $\mathfrak{m}_R^n$ recovers $\mathscr{O}^\wedge_{X_n,x_0}$ due to the relationship between X and the X_n's. But X_n is an actual moduli scheme over the ring $R/\mathfrak{m}_R^n$ (unlike X over R). In view of the assumed universal property of (ρ, R), it follows that $\mathscr{O}^\wedge_{X_n,x_0}$ is the deformation ring for ordinary crystalline structures lifting $\overline{L}_k$ on cyclotomic-determinant deformations of ρ_k (possibly with framing) having coefficients in Λ-finite artin local rings whose nth power vanishes. Hence, $\mathscr{O}^\wedge_{X,x_0}$ is the analogous formal deformation ring for arbitrary Λ-finite artin local coefficients (without restriction on the nilpotence order of the maximal ideal).

Our problem is therefore to prove that there is no obstruction to infinitesimal deformation of $(\rho_k, \overline{L}_k)$ as ordinary crystalline structures with cyclotomic determinant. (There is no obstruction when we impose the additional data of a framing, as that amounts to simply lifting some bases through a surjection of finite free modules.) That is, given such a representation over B/J we wish to lift it to one over B. The given representation with ordinary crystalline structure over B/J has diagonal characters $\{\chi\eta^{-1}, \eta\}$ for some unramified continuous $\eta : G_K \to (B/J)^\times$. Since $G_K/I_K = G_k = \widehat{\mathbf{Z}}$, we can lift η to an unramified continuous $\widetilde{\eta} : G_K \to B^\times$ (choose an arbitrary lift of $\eta(\mathrm{Frob}_k) \in (B/J)^\times$). We claim that this lifts to an ordinary crystalline deformation to B with diagonal characters $\{\chi\widetilde{\eta}^{-1}, \widetilde{\eta}\}$.

Thinking in terms of "upper-right matrix entries", we have to prove the surjectivity of the natural map

$$\mathrm{H}^1_{\mathrm{crys}}(K, B(\chi\widetilde{\eta}^{-2})) \to \mathrm{H}^1_{\mathrm{crys}}(K, (B/J)(\chi\eta^{-2})).$$

For $M = B(\widetilde{\eta}^{-2})$ we have $M/JM = B(\eta^{-2})$, and these are unramified discrete G_K-modules killed by a power of p. By Lemma 3.10, the natural map

$$\mathrm{H}^1_{\mathrm{crys}}(K, M(\chi)) \to \mathrm{H}^1_{\mathrm{crys}}(K, (M/JM)(\chi))$$

is surjective, so we are done. ■

Proposition 4.2. *The natural map $f_E : X_E \to \mathrm{Spec}(R_E)$ is a closed immersion, and this regular closed subscheme meets $\mathrm{MaxSpec}(R_E)$ in precisely the set of closed points $x \in \mathrm{MaxSpec}(R_E)$ such that $\rho_x : G_K \to \mathrm{GL}_2(E(x))$ is ordinary crystalline in the sense of Definition 3.1. In particular, the locus of ordinary crystalline points in $\mathrm{MaxSpec}(R_E)$ is Zariski-closed.*

Proof. By Proposition 2.5, it suffices to prove that f_E is injective on C-valued points for all E-finite C. We may assume C is local, so its residue field E' is E-finite. By Hensel's Lemma and the separability of E'/E, the E-algebra structure on C uniquely extends to an E'-algebra structure lifting the residue field. Thus, $C = E' \oplus I$ for the nilradical I of C. In particular, if Λ' denotes the valuation ring of E' then $\Lambda' \oplus I$ is a local Λ'-subalgebra of C, though it is generally not Λ'-finite since the E'-vector space I is generally nonzero.

To prove that the map $X(C) \to (\mathrm{Spec}\, R)(C)$ between sets of C-valued points over Λ is injective, we can first replace $f : X \to \mathrm{Spec}\, R$ with its scalar extension by $\Lambda \to \Lambda'$. This scalar extension is compatible with the formation of X (as may be checked on infinitesimal fibers over $\mathrm{Spec}\, R$), so we may assume $E' = E$.

We fix a map $\overline{\phi} : \mathrm{Spec}\, C \to \mathrm{Spec}\, R$ over Λ and seek to prove that it has at most one lift to a map $\phi : \mathrm{Spec}\, C \to X$. Consider the Λ-algebra map

$$\Lambda[\![x_1, \ldots, x_m]\!]/(f_1, \ldots, f_s) = R \to C = E \oplus I$$

corresponding to $\overline{\phi}$. Passing to the quotient $C/I = E$, this map carries each x_j to (c_j, y_j) for some $y_j \in I$ and $c_j \in (\pi)$. Thus, we can make the formal change of parameters $x_j \mapsto x_j - c_j$ to get to the case when $x_j \mapsto y_j \in I$ for all j. Since $I^N = 0$ for some large N, any monomial in the x's with large enough degree maps to 0 in C. Hence, the image of R in C is contained in $\widetilde{C} := \Lambda \oplus J$ for a finite Λ-submodule $J \subset I$, and we can increase J so that $\widetilde{C}$ is a local finite flat Λ-subalgebra of C. Note that $\widetilde{C}_E = C$.

Now consider any ϕ lifting $\overline{\phi}$. The restriction of ϕ to the closed point of $\operatorname{Spec} C$ is a map $\phi_0 : \operatorname{Spec} E \to X$ over the specialization $\overline{\phi}_0 : \operatorname{Spec} E \to R$ over Λ. This latter specialization is a Λ-algebra map $R \to E$ and hence lands in Λ. This resulting Λ-valued point of $\operatorname{Spec} R$ *uniquely* lifts to a Λ-valued point of the R-proper X extending ϕ_0, due to the valuative criterion for properness. Consider an open affine $\operatorname{Spec} A$ in X around the image of the Λ-valued point of X extending ϕ_0. This open affine contains the image of ϕ since $E = C/I$ with I nilpotent. The resulting Λ-algebra map

$$A \to C = E \oplus I$$

with A a *finite type* R-algebra lands in $\Lambda \oplus I$ by the choice of A, and so lands in $\widetilde{C}$ upon taking J big enough.

We have now constructed an R-map $\operatorname{Spec} \widetilde{C} \to X$ that serves as an "integral model" (over Λ) for ϕ. The choice of J can always be increased even further, so to prove the uniqueness of ϕ (if one exists) it suffices to consider a pair of R-maps $\operatorname{Spec} \widetilde{C} \rightrightarrows X$ over a *common* local Λ-map $\operatorname{Spec} \widetilde{C} \to R$, and to show that the resulting "generic fiber" maps $\operatorname{Spec} C \rightrightarrows X$ coincide.

By locality of the Λ-map $R \to \widetilde{C}$ to the Λ-finite $\widetilde{C}$, a cofinal system of open ideals in $\widetilde{C}$ is given by the $\mathfrak{m}_R^n \widetilde{C}$. Using the formal GAGA construction of X from the X_n's, it follows that the pair of maps $\operatorname{Spec} \widetilde{C} \rightrightarrows X$ corresponds to a pair of ordinary crystalline structures on $\rho_{\widetilde{C}}$ (i.e., compatible such structures over each artinian quotient of $\widetilde{C}$). These coincide provided that the resulting pair of filtrations on ρ_C coincide, since a $\widetilde{C}$-line in $\widetilde{C}^2$ is uniquely determined by the associated C-line in C^2 (via saturation of Λ-finite submodules in finite-dimensional E-vector spaces, as $\widetilde{C}$ is finite flat over Λ and $\widetilde{C}_E = C$).

Thus, the injectivity of f_E on C-points is reduced to proving that if $G_K \to \mathrm{GL}_2(C)$ is a homomorphism admitting an upper triangular form

$$\begin{pmatrix} \chi\eta^{-1} & * \\ 0 & \eta \end{pmatrix}$$

relative to some C-basis with $\eta : G_K \to C^\times$ unramified then the $\chi\eta^{-1}$-line is uniquely determined. This C-line is precisely the locus of vectors on which I_K acts through χ. Indeed, to prove this it suffices to check that in the quotient C-line η the space of χ-isotypic vectors for the I_K-action vanishes. Since χ is valued in $E^\times$ and C has an E-linear filtration by ideals with successive codimension 1 over E, we just need to observe that $\chi \neq 1$ in $E^\times$ since $\mathrm{char}(E) = 0$. (In more sophisticated p-adic Hodge theory settings, the analogue of this step requires results such as Tate's isogeny theorem for p-divisible groups over Λ: uniqueness results for integral structures in case of generic characteristic 0.)

To identify the closed points of X_E within $\mathrm{MaxSpec}(R_E)$, we will use Example 3.9. Closed points of X_E are obtained from E'-valued points for finite extension fields E'/E; let Λ' be the valuation ring of E'. The preceding argument with the valuative criterion for properness shows that any E'-valued point of X_E uniquely extends to a Λ'-valued point of X over $\operatorname{Spec} R$, and such a point corresponds precisely to a filtration on $\rho_{E'}$ as in Definition 3.1, due to Example 3.9. This proves that the closed points of X_E are the ordinary crystalline points of $\mathrm{MaxSpec}(R_E)$. ■

Corollary 4.3. *Let $\rho_0 : G_K \to \mathrm{GL}_2(k)$ be an ordinary crystalline representation with cyclotomic determinant. Let $\rho : G_K \to \mathrm{GL}_2(R)$ be the universal framed deformation with cyclotomic determinant; if ρ_0 has only scalar endomorphisms we allow alternatively that (ρ, R) is the universal deformation of ρ_0 with cyclotomic determinant.*

The locus of ordinary crystalline points in $\mathrm{MaxSpec}(R_E)$ *is Zariski-closed, and if*

$$\mathrm{Spec}(R^{\mathrm{ord}})$$

denotes the Zariski closure in $\mathrm{Spec}(R)$ *of this locus in* $\mathrm{Spec}(R_E)$ *then R_E^{ord} is regular, and it is a domain except precisely when $(\rho_0)_{\overline{k}} = \psi_1 \oplus \psi_2$ with $\psi_1 \neq \psi_2$ and each ψ_i an unramified $\overline{k}^\times$-valued twist of $\omega := \chi \bmod p$.*

Note that the exceptional cases at the end of the corollary do not include the case when ρ_0 makes G_K act trivially. This is really useful: we will apply this corollary later for the universal framed deformation ring of a 2-dimensional *trivial* residual representation (after making a preliminary finite extension on K).

Proof. Apply the preceding theory to ρ, so we get the "moduli scheme" $f : X \to \mathrm{Spec}\, R$ that is regular and induces a closed immersion f_E over E whose image on closed points is the set of ordinary crystalline points of $\mathrm{MaxSpec}(R_E)$. This gives the Zariski-closedness and regularity claims, so the domain property amounts to the assertion that X_E is connected. We saw above that $X \bmod \pi$ is reduced, so by Lemma 2.2 the connectedness of X_E is equivalent to the connectedness of the proper special fiber $f_0 : X_0 \to \mathrm{Spec}\, k$ (since X_0 is certainly non-empty, due to its moduli-theoretic meaning and the fact that the ordinary crystalline hypothesis on ρ_0 provides a k-point of X_0).

Loosely speaking, X_0 is the moduli scheme of ordinary crystalline structures on ρ_k. That is, it parameterizes all G_K-stable lines in ρ_0 on which I_K acts by χ. By construction, the non-empty X_0 is a closed subscheme of $\mathbf{P}^1_k$, so it is connected except precisely when it is not the entire projective line nor is a single geometric point (as we know $X_0(k) \neq \emptyset$). The condition $X_0 = \mathbf{P}^1_k$ says that every line in $(\rho_0)_{\overline{k}}$ is G_K-stable and has I_K-action by ω. In other words, ρ_0 is a scalar representation via an unramified twist of ω. Thus, the disconnectedness case is when $(\rho_0)_{\overline{k}}$ has more than one – but only finitely many! – G_K-stable line with action by an unramified twist of ω. The number of such lines is therefore exactly two, by the Jordan-Hölder theorem. Such cases are precisely when $(\rho_0)_{\overline{k}}$ is a direct sum of distinct $\overline{k}^\times$-valued characters of G_K which are each an unramified twist of ω. ∎

Remark 4.4. There is a variant of the preceding considerations which is useful in practice: require the determinant of ρ_0 and its deformations to be $\chi\psi$ for a fixed unramified (possibly nontrivial) continuous character $\psi : G_K \to \Lambda^\times$.

In such cases the conclusions of Corollary 4.3 hold, by essentially the same proof. The point is that to prove these claims we can first make a scalar extension from Λ to the valuation ring of a finite extension of E, so we can arrange that the unramified $\psi : G_K \to \Lambda^\times$ admits a square root. It is then harmless in the Galois deformation theory to twist everything by the reciprocal of this square root, so we are thereby reduced back to the case $\psi = 1$ which was treated above.

The final application we take up is the determination of the dimension of the regular Zariski closure X_E of the locus of ordinary crystalline points in $\mathrm{Spec}(R_E)$ for a framed deformation ring $R_{\rho_0}^{\square,\det=\chi\psi}$. In some cases this Zariski closure is disconnected, but we claim that its connected components are always of the same pure dimension:

Proposition 4.5. *The E-fiber of the ordinary crystalline framed deformation ring*

$$R_{\rho_0}^{\square,\mathrm{ord},\det=\chi\psi}$$

has dimension $3+[K:\mathbf{Q}_p]$ at all closed points.

Proof. As above, we may reduce to the case $\psi = 1$ by making a suitable finite extension on E (which is harmless for our purposes). In view of the regularity, we just have to compute the dimension of the *tangent space* at each closed point in characteristic 0. This will be a Galois H^1 with coefficients in a p-adic field.

By the proof of Proposition 4.2 (relating C-valued points and $\widetilde{C}$-valued points) and the lecture in the fall on charateristic-0 deformation rings, if we identify a closed point $x \in X_E$ with an ordinary crystalline representation $\rho_x : G_K \to \mathrm{GL}_2(E(x))$ then the (regular) completed local ring of X_E at x is the deformation ring of ρ_x relative to the conditions of having determinant χ and being ordinary crystalline (in the sense of Definition 3.1, generalized in the evident manner to allow coefficients in any finite E-algebra, not just finite extension fields of E).

The method of the proof of the Corollary in §1 of Samit's lecture in the fall shows (OK, this should be revised for clarity!) carries over to characteristic-zero deformation theory, so the dimension of the tangent space to the cyclotomic-determinant framed deformation functor exceeds the dimension of the tangent space to the cyclotomic-determinant deformation functor by $\dim \mathrm{PGL}_2 + h^0(\mathrm{ad}^0(\rho_x)) = 3 + h^0(\mathrm{ad}^0(\rho_x))$. Thus, the problem is to prove that in the tangent space $\mathrm{H}^1(K, \mathrm{ad}^0(\rho_x))$ to the cyclotomic-determinant deformation ring of ρ_x, the space of first-order deformations which are ordinary crystalline has $E(x)$-dimension $[K:\mathbf{Q}_p]$ if the reducible ρ_x has only scalar endomorphisms and $1+[K:\mathbf{Q}_p]$ otherwise (the case when ρ_x is a direct sum of characters, necessarily distinct due to ramification considerations). The representation ρ_x has the form

$$\rho_x \simeq \begin{pmatrix} \chi\eta^{-1} & * \\ 0 & \eta \end{pmatrix}$$

for some unramified $\eta : G_K \to \mathscr{O}_{E(x)}^\times$, and upon restriction to I_K (which kills η and η^{-1}) the resulting class in $\mathrm{H}^1(I_K, \mathbf{Z}_p(1))$ arises from units in $\mathscr{O}_{K^{\mathrm{un}}}^\wedge$ via Kummer theory (i.e., it is killed by the natural map $\mathrm{H}^1(I_K, \mathbf{Z}_p(1)) \to \mathbf{Z}_p$ defined by the valuation map $K^{\mathrm{un}\times} \to \mathbf{Z}$ and Kummer theory). It is harmless to rename $E(x)$ as E, so this is now a very concrete problem in Galois cohomology and Kummer theory using the "explicit" upper-triangular description of ρ_x.

The only method we know to carry out the dimension calculation in the general case is to bring in deeper methods related to p-adic Hodge theory or p-divisible groups over very ramified p-adic discrete valuation rings. But we will only apply the Proposition in the special case that ρ_0 is the *trivial* 2-dimensional residual representation. So now we will give a proof only in this case. Note that the residual triviality forces the mod-p cyclotomic character of G_K to be trivial, so there is a distinguished ordinary crystalline lift with cyclotomic determinant:

$\rho_\chi := E(\chi) \oplus E$. We have seen that for our ρ_0, the ord-crystalline framed deformation ring R with determinant χ has the property that R_E is regular with *connected* spectrum. Provided that the $E(x)$-dimension of its tangent space at any closed point $x \in \mathrm{MaxSpec}(R_E)$ is *independent* of x, it would suffice to carry out the dimension computation at a single x. For example, we would be reduced to computing the E-dimension of the ord-crystalline subspace of

$$\mathrm{H}^1(K, \mathrm{ad}^0(\rho_\chi)) = \mathrm{Ext}^1_K(\rho_\chi, \rho_\chi)^{\det=\chi}.$$

The equidimensionality of the tangent spaces on $\mathrm{MaxSpec}(R_E)$ is a special case of:

Lemma 4.6. *Let R be a quotient of a formal power series ring over Λ, and assume that R_E is normal with connected spectrum. Then all maximal ideals of R_E have the same height.*

Proof. Replacing R with the quotient by its nilradical has no effect on R_E, so we can assume that R is reduced. Likewise we can assume it has vanishing π-power torsion, so R is Λ-flat. Hence, R is a domain (as R_E is a domain, due to regularity and connectedness of its spectrum). But R is excellent, so the normalization map $R \to R'$ is module-finite. The residue field may increase in the normalization process, so R' is a quotient of a formal power series ring over the valuation ring of some finite unramified extension E' of E. Then $R_E = R'_E = R'_{E'}$, so we can replace (R, E) with (R', E') to reduce to the case when R is a normal domain.

There are now two ways to proceed: commutative algebra, or rigid-analytic geometry. For the commutative algebra method, let $\mathfrak{p}$ be a maximal ideal of $R_E = R[1/p]$. The complete local noetherian domain R is catenary ([5, 31.6(iv)]; in general, the catenary property is also part of the definition of excellence), so $\dim R_P + \dim(R/P) = \dim(R)$ for any prime ideal P of R. Taking P corresponding to $\mathfrak{p}$, we get

$$\dim(R_E)_\mathfrak{p} = \dim(R_P) = \dim(R) - \dim(R/P),$$

so it suffices to prove that $\dim(R/P) = 1$ for all such P. The quotient R/P is a Λ-flat quotient of a formal power series ring over Λ such that its generic fiber ring is $R_E/\mathfrak{p}$, which is a field of finite degree over E. Hence, the *subring* R/P lies in the valuation ring of this finite extension of E, whence R/P is module-finite over Λ and so is of dimension 1. This completes the commutative algebra proof.

We merely sketch the rigid-analytic method, which provides nice geometric intuition (and can be made rigorous). By choosing a presentation

$$R \simeq \Lambda[\![t_1, \ldots, t_m]\!]/(f_1, \ldots, f_s),$$

it is natural to associate to R the rigid-analytic space M over E defined by $f_1 = \cdots = f_s = 0$ in the open unit m-disc over E. This construction is given in more intrinsic terms in [1, 7.1]. That exposition proves some very useful related facts: there is a natural bijective correspondence between $\mathrm{MaxSpec}(R_E)$ and the underlying set of M such that the completed local rings at corresponding points are naturally E-isomorphic [1, 7.1.9], and R_E is identified with the ring of bounded global analytic functions on M (here we use the normality of R) [1, 7.3.6]. Thus, the completed local rings on M are normal (as the excellent R_E is normal, by hypothesis, so its completed local rings are normal). But it is elementary to check that a noetherian local ring is a normal domain if its completion is, so the analytic local rings

on M are normal and the affinoid opens in M have normal coordinate ring. Moreover, M is *connected* for the Tate topology since idempotents are bounded analytic functions and R_E is a domain (due to normality and connectedness of its spectrum).

Since the completed local rings of R_E at its maximal ideals coincide with the completed local rings on M, and completion preserves dimension for local noetherian rings, to prove that all maximal ideals of R_E have the same height it suffices to prove that all local rings on M have the same dimension. More generally, any normal rigid-analytic space has pointwise dimension that is locally constant for the Tate topology (and hence globally constant in the *connected* case): this comes down to the fact that an affinoid space associated to a domain has constant pointwise dimension, which is [2, Lemma 2.1.5].

As an alternative argument in the rigid-analytic case if we assume R_E is regular (as holds in the cases we need), regularity of M implies smoothness of M since $\mathrm{char}(E) = 0$, so the coherent sheaf $\Omega^1_{M/E}$ is locally free on M with rank $\dim_m(M)$ at any $m \in M$. But connectedness of M forces this rank to be globally constant, whence M has constant pointwise dimension as desired. ■

Returning to the proof of Proposition 4.5, we just have to prove that the ordinary crystalline subspace of $\mathrm{Ext}^1_K(\rho_\chi, \rho_\chi)^{\det=\chi}$ has E-dimension equal to $[K : \mathbf{Q}_p]$. Since $\rho_\chi = E(\chi) \oplus E$, we have an equality of E-vector spaces

$$\mathrm{Ext}^1_K(\rho_\chi, \rho_\chi) = \mathrm{Ext}^1_K(E(\chi), E(\chi)) \oplus \mathrm{Ext}^1_K(E(\chi), E) \oplus \mathrm{Ext}^1(E, E(\chi)) \oplus \mathrm{Ext}^1_K(E, E). \quad (4.1)$$

The condition that an $E[\epsilon]$-deformation of ρ_χ has cyclotomic determinant amounts to the condition that its Ext-class ξ on the left side of (4.1) has components in outer terms that are Cartier dual to each other (as one checks with a direct 4×4 matrix calculation). The ordinarity condition likewise amounts to the vanishing of the $\mathrm{Ext}^1_K(E(\chi), E)$ component. The crystalline condition then says that the component in $\mathrm{Ext}^1_K(E, E) = \mathrm{H}^1(K, E)$ is unramified (a 1-dimensional subspace) and the component in $\mathrm{Ext}^1_K(E, E(\chi)) = \mathrm{H}^1(K, E(1))$ is crystalline.

Since $\rho_\chi = E(\chi) \oplus E$ has non-scalar endomorphisms, our problem is to prove that the ordinary-crystalline Ext-space with cyclotomic determinant inside of the left side of (4.1) has E-dimension $1 + [K : \mathbf{Q}_p]$. We have already accounted for one dimension, and it remains to prove that $\mathrm{H}^1_{\mathrm{crys}}(K, E(1)) = \mathrm{H}^1_{\mathrm{crys}}(K, \mathbf{Z}_p(1)) \otimes_{\mathbf{Z}_p} E$ has E-dimension $[K : \mathbf{Q}_p]$. But $\mathrm{H}^1_{\mathrm{crys}}(K, \mathbf{Z}_p(1))$ is the multiplicative p-adic completion of $\mathscr{O}_K^\times = k^\times \times (1 + \mathfrak{m}_K)$, which is $1 + \mathfrak{m}_K$. Via the logarithm, its $\mathbf{Z}_p$-rank as a multiplicative $\mathbf{Z}_p$-module is $[K : \mathbf{Q}_p]$. ■

References

[1] A.J. de Jong, "Crystalline Dieudonné theory via formal and rigid geometry", Publ. Math. IHES **82** (1995), pp. 5–96.

[2] B. Conrad, "Irreducible components of rigid spaces", Annales Fourier **49**(2), 1999, pp. 473–541.

[3] M. Kisin, "Modularity of 2-adic Barsotti-Tate representations", Inv. Math. **178**(3), 2009, pp. 587–634.

[4] H. Matsumura, *Commutative algebra* (2nd ed.), Benjamin, 1980.

[5] H. Matsumura, *Commutative ring theory*, Cambridge Advanced Studies in Mathematics **8**, Cambridge, 1986.

LOCAL PROPERTIES OF MODULAR GALOIS REPRESENTATIONS

ANDREW SNOWDEN

1. Introduction

Let f be a cuspidal eigenform of weight 2 and level $\Gamma_0(N)$. Let p be a prime, which we assume does not divide N. We have stated (though have not proved) that there exists a Galois representation $\rho : G_{\mathbf{Q},S} \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$, where S is the set of primes dividing pN, which satisfies and is characterized by the following two properties: (1) the determinant of ρ is the cyclotomic character $\chi = \chi_p$; and (2) for a prime $\ell \nmid pN$ the trace of $\rho(\mathrm{Frob}_\ell)$ is equal to the eigenvalue of the Hecke operator T_ℓ acting on f. We have also stated (and not proved) that for $\ell \mid N$ the representation $\rho|_{G_{\mathbf{Q}_\ell}}$ corresponds under local Langlands to the local component of the automorphic representation of f at ℓ. We have not yet examined the local representation $\rho|_{G_{\mathbf{Q}_p}}$. For the purposes of this seminar, we will need only one result: if f is ordinary (in the sense of modular forms) then $\rho|_{G_{\mathbf{Q}_p}}$ is crystalline and ordinary (in the sense of Galois representations). The definitions of ordinary are recalled below.

The purpose of this lecture is to sketch the construction of ρ and the proofs that it satisfies the above local conditions, at least for $\ell \nmid N$. The representation ρ is found as a quotient of the Jacobian $J_0(N)$ of the modular curve $X_0(N)$, and is not difficult to construct. To establish the properties of ρ at the unramified places and at p, we use the Eichler-Shimura relation. To formulate and prove this identity, we use the reduction of $X_0(N)$ modulo p. This requires us to introduce some of the theory of moduli of elliptic curves over integers; fortunately, we are in a rather easy situation. At the end of these notes, we explain a bit about what happens in the Hilbert modular case.

I should say here that I am not extremely familiar with this material. I believe I have the main points correct, but I might have some details wrong. Certainly, some details have been omitted. For certain topics, more complete treatments can be found in the references.

2. Moduli of elliptic curves

In this section we define moduli spaces of elliptic curves and establish some of their basic properties.

2.1. **The moduli space.** Let S be a scheme. An *elliptic curve* over S is a smooth proper group scheme $E \to S$ whose geometric fibers are connected genus 1 curves. Let Y be the functor which assigns to a scheme S the groupoid $Y(S)$ of elliptic curves over S; that is, $Y(S)$ is the category whose objects are elliptic curves over S and where morphisms are isomorphisms of group schemes over S. We call Y the *moduli space of elliptic curves*.

Proposition 1. *The functor Y is a stack.*

Proof. Let E/S be an elliptic curve. The zero section $0 : S \to E$ defines an ample divisor D on E (in the relative sense) and $3D$ is very ample. Let A_E be the projective coordinate ring of E in this embedding, that is, A_E is $\bigoplus_{n \geq 0} f_*(\mathscr{O}(3nD))$ where $f : E \to S$ is the structure map. Then A_E is a quasi-coherent sheaf of graded rings on S and E is identified with $\mathrm{Proj}(A_E)$. The functor $E \mapsto A_E$ identifies $Y(S)$ with a subcategory of the category of quasi-coherent graded algebras on S. The latter forms a stack in the fppf topology by Grothendieck's theory of flat descent. It is easy to conclude from this that Y itself forms a stack. More precisely, let E_i be elliptic curves on a cover U_i of a scheme S and let f_{ij} be an isomorphism of E_i and E_j on U_{ij} satisfying the 1-cocycle condition. Then A_{E_i} and $A_{f_{ij}}$ define descent data for algebras on S. By flat descent, one obtains a quasi-coherent graded algebra A on S. Put $E = \mathrm{Proj}(S)$. One has a canonical identification $E|_{U_i} = E_i$, which allows one to establish the geometric properties required of E, as these properties are fppf local. (One must say a bit more along the same lines to get the zero section and group law on E.) □

Date: September 11, 2010.

Proposition 2. *The functor Y is formally smooth (over $\mathbf{Z}$).*

Proof. Let S be an affine scheme and let S_0 be a closed subscheme defined by a square zero ideal I. Let E_0 be an elliptic curve over S_0. We must extend E_0 to an elliptic curve over S. Note that any such curve will have the same underlying topological space as E, just a different structure sheaf. Let $U_{0,i}$ be an open affine cover of E_0. Since smooth affine schemes always lift, we can find a smooth affine U_i over S extending $U_{0,i}$. We have thus extended $\mathscr{O}_{E_0}$ to an $\mathscr{O}_S$-algebra on an open cover. These bigger algebras may not patch together, but we can try to modify them so that they do. There is an obstruction class in $H^2(E_0, T_{E_0})$ measuring if such a modification is possible; here T_{E_0} is the tangent sheaf of E_0. Now, since S_0 is affine, this cohomology group is equal to $H^0(S_0, R^2 f_* T_{E_0})$, where $f : E_0 \to S_0$ is the structure map. As $E_0 \to S_0$ is a curve, $R^2 f_*$ vanishes. This shows that $H^2(E_0, T_{E_0})$ vanishes as well, and thus there is no obstruction to the modification procedure. We have thus found a smooth scheme E over S such that E_{S_0} is canonically identified with E_0. One then needs to extend the zero section from S_0 to S; we leave this to the reader. □

2.2. **Level structure in good characteristic.** Let N be an integer. For an elliptic curve E/S we write $E[N]$ for the N-torsion of E. It is a finite flat group scheme over S. If N is invertible on S then $E[N]$ is a finite étale group scheme over S. We define three additional moduli spaces $Y(N)$, $Y_1(N)$ and $Y_0(N)$ over $\mathbf{Z}[1/N]$, as follows:

- $Y(N)(S)$ is the category of tuples (E, P, Q) where E/S is an elliptic curve $P, Q \in E[N]$ form a basis of $E[N]$, i.e., the map $((\mathbf{Z}/N\mathbf{Z})^2)_S \to E[N]$ defined by (P, Q) is an isomorphism of sheaves.
- $Y_1(N)(S)$ is the category of pairs (E, P) where E/S is an elliptic curve and $P \in E[N]$ is a point of exact order N, i.e., the map $(\mathbf{Z}/N\mathbf{Z})_S \to E$ defined by P is an injection of sheaves.
- $Y_0(N)(S)$ is the category of pairs (E, G) where E/S is an elliptic curve and $G \subset E[N]$ is a subgroup scheme which is fppf locally isomorphic to $(\mathbf{Z}/N\mathbf{Z})_S$.

For $N \geq 3$ the category $Y(N)$ is discrete; the same holds for $Y_1(N)$ and $Y_0(N)$ for N large enough. We assume from now on that N is sufficiently large for this to be the case. We now have the following result:

Proposition 3. *Each of $Y(N)$, $Y_1(N)$ and $Y_0(N)$ is a smooth affine curve over $\mathbf{Z}[1/N]$. The natural map from each to Y is finite and étale.*

Proof. We consider only $Y(N)$, leaving the others to the reader. First, it follows easily from Proposition 1 that $Y(N)$ is itself a stack; it is therefore a sheaf of sets since it is discrete. We now show that $Y(N) \to Y$ is relatively representable, finite and étale. Let $S \to Y$ be a map, corresponding to an elliptic curve E/S. The fiber product $Y(N) \times_Y S$ is then identified with the subsheaf of $E[N] \times E[N]$ consisting of those pairs of sections which form a basis. This is clearly a finite étale scheme over S. This establishes the claim. The formal smoothness of $Y(N)$ now follows from Proposition 2.

Here is the main idea of one approach to get representability. The group $G = \mathrm{GL}(2, \mathbf{Z}/3\mathbf{Z})$ acts on $Y(3)$. One can write down explicit equations for $Y(3)$ demonstrating that it is a smooth affine curve over $\mathbf{Z}[1/3]$. One can also show that G acts freely on $Y(N) \times_Y Y(3)$ and that the quotient is identified with $Y(N)$. Since $Y(N)$ is relatively representable and finite étale, the product $Y(N) \times_Y Y(3)$ is a smooth affine curve over $\mathbf{Z}[1/3N]$. It follows that the same holds for the quotient by G, which establishes the required properties of $Y(N)$, at least over $\mathbf{Z}[1/3N]$. There is another explicit moduli problem and finite group one can use to obtain the required properties over $\mathbf{Z}[1/2N]$. This implies the results over $\mathbf{Z}[1/N]$. (One can probably avoid the use of $Y(3)$ by appealing to more general results, such as Artin's representability theorem.) □

Remark 4. The finite étale covers of Y provided by $Y(N)$ show that Y is a Deligne-Mumford stack. The same is true for $Y(N)$, $Y_0(N)$ and $Y_1(N)$ when N is small.

2.3. **Compactification.** The modular curve Y constructed in the previous section is affine. We would now like to compactify it. To do this we must add a few points to it. These points correspond to curves which are limits of elliptic curves. To see what a "limit of an elliptic curve" is, it is useful to think about the situation over the complex numbers: to degenerate an elliptic curve, one can take a few cycles on it and pinch them each to a point. The result is a bunch of $\mathbf{P}^1$s glued together. This motivates the formal definitions which follow.

Let n be an integer. Let C be the scheme obtained by taking $\mathbf{P}^1 \times \mathbf{Z}/n\mathbf{Z}$ and identifying the point 0 in the ith $\mathbf{P}^1$ with the point ∞ in the $(i+1)$st $\mathbf{P}^1$. We call C an *n-gon*. We let C° denote the smooth part of C. The space C° is identified with $\mathbf{G}_m \times \mathbf{Z}/n\mathbf{Z}$, and is thus naturally a group. Furthermore, the group law on C° extends to an action on all of C.

A *generalized elliptic curve* over a scheme S is a proper flat curve $E \to S$ together with a multiplication map $E^\circ \times E \to E$ which gives E° the structure of a group scheme, in such a way that the fibers of E are elliptic curves or polygons (respecting the obvious structure). Here E° denotes the open subset of E where the fibers are smooth. We define $X(S)$ to be the groupoid of generalized elliptic curves E/S whose fibers are all irreducible, i.e., elliptic curves or 1-gons.

Proposition 5. *The functor X is a proper smooth Deligne-Mumford stack over $\mathbf{Z}$.*

Proof. This is proved in [DR] using Artin's representability theorem. I imagine one could give an argument similar to the one we gave for Y. Properness can be seen from the valuative criterion. Let A be a valuation ring with fraction field K and let E/K be an elliptic curve. The semi-stable reduction theorem implies that there is a finite extension K'/K such that the base change E' of E to K' has good or multiplicative reduction — that is, its minimal Weierstrass equation defines a scheme over A' having semi-stable reduction. This shows that the point of $X(K)$ coming from E, when mapping into $X(K')$, comes from an element of $X(A')$. Thus X satisfies the valuative criterion for properness. (Note that this criterion is a little bit different than the one for schemes: we are allowed to extend the field K.) □

We can also compactify the spaces $Y(N)$, $Y_0(N)$ and $Y_1(N)$ over $\mathbf{Z}[1/N]$. To do this, we need to define the notion of a level structure on a generalized elliptic curve. Thus let E/S be a generalized elliptic curve. We let $E[N]$ be the N-torsion of the group E°. The only subtlety concerning level structures is that we require them to be ample, which amounts to them meeting every irreducible component of the fibers of E. Thus a $\Gamma_0(N)$ structure is a subgroup $G \subset E[N]$ which is locally isomorphic to $\mathbf{Z}/N\mathbf{Z}$ and such that G meets each irreducible component of the fibers of E. Note that this imposes a restriction on what the fibers can be: their component group must be a quotient of $\mathbf{Z}/N\mathbf{Z}$. We define $X_0(N)$ to be pairs (E, G) where E is a generalized elliptic curve whose fibers are elliptic curves or N-gons and $G \subset E[N]$ is a $\Gamma_0(N)$ structure, as defined above. The spaces $X(N)$ and $X_1(N)$ are defined similarly.

Proposition 6. *The functors $X(N)$, $X_0(N)$ and $X_1(N)$ are smooth proper schemes over $\mathbf{Z}[1/N]$ (assuming N large enough).*

2.4. **The space $X_0(p)$ over Z.** The compactified space $X_0(N)$ — and indeed, even the open curve $Y_0(N)$ — has only been defined over $\mathbf{Z}[1/N]$. It is a bit tricky to formulate what these spaces should be over $\mathbf{Z}$ since one has to specify what it means for a (non étale) group scheme to be cyclic. However, when N is a prime, this is not hard: every group of order p should be considered cyclic! We thus define $Y_0(p)(S)$ (resp. $X_0(p)(S)$) to be the groupoid of pairs (E, G) where E is an elliptic curve (resp. generalized elliptic curve) over S and $G \subset E[p]$ is a finite flat subgroup scheme of order p which is ample (this condition is only relevant for $X_0(p)$). We then have the following result:

Proposition 7. *The functor $X_0(p)$ is a proper flat curve over $\mathbf{Z}$ (for p large).*

We will actually need to extend this a bit for our application. Let N be an integer prime to p. A cyclic group of order Np decomposes canonically as a product of a cyclic group of order N and one of order p. We thus have $X_0(Np) = X(N) \times_X X(p)$ over $\mathbf{Z}[1/Np]$. We take this formula as the *definition* of $X_0(Np)$ over $\mathbf{Z}[1/N]$. That is, $X_0(Np)$ consists of tuples (E, G, H) where E/S is a generalized elliptic curve whose fibers are elliptic curves or pN-gons, $G \subset E[N]$ is a group locally isomorphic to $\mathbf{Z}/NZ$ and $H \subset E[p]$ is a finite flat subgroup such that GH meets every irreducible component of the fibers of E. We then have:

Proposition 8. *The functor $X_0(pN)$ is a proper flat curve over $\mathbf{Z}[1/N]$ (for Np large).*

3. Elliptic curves in characteristic p

We now examine elliptic curves in characteristic p and their moduli. We establish the Eichler-Shimura relation.

3.1. **Group schemes.** Let k be a finite field. Let G/k be a finite commutative group scheme. We say that G is *local* if it is connected. There is a canonical exact sequence

$$1 \to G^\circ \to G \to G^{\text{et}} \to 1$$

where G° is local and G^{et} is étale. If the order of G is prime to p then G is automatically étale.

Define a functor $G^\vee$ by $G^\vee(T) = \mathrm{Hom}(G_T, (\mathbf{G}_m)_T)$. Then $G^\vee$ is again a finite group scheme over k. We call $G^\vee$ the *Cartier dual* of G. Cartier duality is an anti-equivalence of categories. The properties "local" and "étale" interact in an interesting manner with Cartier duality. First of all, if G has order prime to p then $G^\vee$ does as well and both are étale. However, the dual of p-power étale group is never étale, and is always connected: for example, $(\mathbf{Z}/p\mathbf{Z})^\vee = \mu_p$. The converse to this is *not* true: for example, if α_p is the kernel of Frobenius on $\mathbf{G}_a$, i.e., $\mathrm{Spec}(k[x]/x^p)$, then α_p is connected and self-dual. We thus find that we can define four classes of groups: étale-étale, étale-local, local-étale and local-local depending on the properties of G and $G^\vee$. Here are examples from the respective classes: $\mathbf{Z}/N\mathbf{Z}$ with N prime to p, $\mathbf{Z}/p\mathbf{Z}$, μ_p and α_p. Every group canonically decomposes as a sum of four groups of these types, so in many circumstances, it suffices to consider groups of only one type. Each type of group forms an abelian category, and with the exception of the étale-étale case, each has only one simple object over $\overline{k}$ (the examples we have listed).

Let G/k be a group scheme. We then have the relative Frobenius map $F : G^F \to G$, which is a map of groups. Here G^F is the Frobenius twist of G; note that $(G^F)^\vee = (G^\vee)^F$. The Cartier dual of the relative Frobenius map is a map $G^\vee \to (G^\vee)^F$. This is not the Frobenius map on $G^\vee$ (it goes in the wrong direction), but a new map, called *Verschebung*, and denoted V. Precisely, this is the Verschebung for $G^\vee$. The Verschebung for G is defined by taking the Cartier dual of the Frobenius map on $G^\vee$; it is a map $V : G \to G^F$. Here are some examples, with $k = \mathbf{F}_p$ (note then that $G^F = G$ for any G). On $\mathbf{Z}/p\mathbf{Z}$ the Frobenius is the identity. On μ_p and α_p the Frobenius is the zero map; these groups are by definition the kernel of Frobenius on $\mathbf{G}_m$ and $\mathbf{G}_a$. Since $\mathbf{Z}/p\mathbf{Z}$ and μ_p are Cartier dual, it follows that $V = 0$ on $\mathbf{Z}/p\mathbf{Z}$ while V is the identity on μ_p. As α_p is self-dual, $V = 0$ on it. Clearly, on the étale-étale groups, F and V are both the identity.

The Frobenius and Vershebung maps in fact allows us to determine which of the four types G is. For instance, G is local if and only if F is nilpotent (meaning $F^n : G^{F^n} \to G$ is zero for $n \gg 0$) and étale if and only if F is an isomorphism. Thus $G^\vee$ is local is and only if V is nilpotent on G and étale if and only if V is an isomorphism on G.

Let A be an abelian variety over k. Then the p-torsion $A[p]$ is an example of a finite group. The Frobenius map on $A[p]$ is nothing other than the Frobenius map on A restricted to $A[p]$. This Frobenius map $F : A^F \to A$ is an isogeny, and thus has a dual $V : A^\vee \to (A^F)^\vee$. We can thus define a map $V : A \to A^F$ by taking the dual to Frobenius on $A^\vee$. The map induced on $A^\vee[p]$ by V is in fact the dual of the Frobenius map on $A[p]$ since Cartier duality and abelian variety duality interact nicely. This is useful when dealing with the torsion groups of abelian varieties, as we will below.

3.2. **Elliptic curves.** Let E be an elliptic curve over k. The group scheme $E[p]$ is finite of order p^2. The Weil pairing

$$E[p] \times E[p] \to \mu_p \subset \mathbf{G}_m$$

implies that $E[p]$ is its own Cartier dual. This implies that there are two possibilities for E: either it is a sum of a local-étale group and an étale-local group each of order p and dual to each other, or else it is local-local. In the first case, $E[p]$ has $\overline{k}$ points while in the second case it does not. We call E *ordinary* in the first case and *supersingular* in the second.

The group scheme $E[p]$ can be exactly determined over $\overline{k}$. In the ordinary case, the étale quotient of $E[p]$ is isomorphic to $\mathbf{Z}/p\mathbf{Z}$ and so the local part, its dual, must be μ_p. Thus $E[p] = \mu_p \oplus \mathbf{Z}/p\mathbf{Z}$. In the supersingular case, $E[p]$ is local-local, and so it is an extension

$$0 \to \alpha_p \to E[p] \to \alpha_p \to 0.$$

There are four such (total spaces of) extensions up to isomorphism: the direct sum (on which $F = 0$ and $V = 0$), the kernel of Frobenius on the Witt scheme W_2 (on which $F = 0$ and $V \neq 0$), the kernel of the square of Frobenius on $\mathbf{G}_a$, namely α_{p^2} (on which $F \neq 0$ and $V = 0$; it is Cariter dual to W_2), and one other (which can be described as the Yoneda sum of the other two non-trivial extensions and on which F and V are each non-zero). The group $E[p]$ is the last one: since F and V on $E[p]$ are the restriction of degree p isogenies on E, their kernels must be order p and therefore cannot be all of $E[p]$; thus F and V are each non-zero on $E[p]$.

3.3. **The space $X_0(p)$.** We now consider $X_0(p)$ over k. This is the space of generalized elliptic curves together with a subgroup of order p (satisfying some condition in the generalized case, which we will ignore). Let E/k be an elliptic curve. How many subgroups of order p does it have? If E is ordinary, then it has

two: the local one and the étale one. It cannot have more than these two, because they are distinct simple objects. If E is supersingular, then it has at most one subgroup of order p, since over $\overline{k}$ it is a non-trivial extension of two simples.

The above discussion shows that we can define two maps $i_1, i_2 : X \to X_0(p)$, by letting $i_1(E)$ be $(E^F, \ker F)$ and $i_2(E)$ be $(E, \ker V)$. It is clear that each map is injective; in fact, each is a closed immersion. Furthermore, the previous discussion shows that they are jointly surjective. To be clear, say that (E, G) is a point of $X_0(p)$. If E is supersingular, then G must coincide with $\ker V$, since there is only one subgroup of order p, and so $(E, G) = i_2(E)$. If E is ordinary and G is étale, then $(E, G) = i_2(E)$, while if E is ordinary and G is local then $(E, G) = i_1(E')$, where E' is such that $(E')^F = E$. Note that if E is ordinary then $i_1(E)$ and $i_2(E)$ are unequal, since in $i_1(E)$ the level structure is étale while in $i_2(E)$ it is local. If E is supersingular so is E^F and thus $i_1(E) = i_2(E^F)$. We therefore find that $X_0(p)$ can be described as two copies of X glued along their supersingular loci identified via $(-)^F$ (at the supersingular points there are nodal singularities). The same discussion applies to $X_0(Np)$: it is two copies of $X_0(N)$ glued along supersingular points by $(-)^F$.

As a side comment, we note that this result shows that the genus of $X_0(p)$ is one less than the number of supersingular curves.

3.4. **Correspondences.** We quickly review the basics of correspondences on curves. Let X be a regular curve over a field k. A *correspondence* on X is a pair $f = (f_1, f_2)$ where f_1 and f_2 are maps from some reduced curve C (the total space) to X such that f_1 is finite. Given two correspondences f and f' with total spaces C and C' we define their *sum*, denoted $f + f'$, to be the natural correspondence with total space $C \amalg C'$. Given a correspondence f with total space C, we get a natural correspondence $\widetilde{f}$ with total space $\widetilde{C}$, the normalization of $\widetilde{C}$, coming from the finite map $\widetilde{C} \to C$. We consider f and $\widetilde{f}$ to be equivalent; note that $\widetilde{C}$ is always regular.

Let $f : X \to Y$ be a map of regular curves over k. We then have a map $f_* : \operatorname{Div}(X) \to \operatorname{Div}(Y)$. This map is characterized by the following two properties: $\deg f_*(D) = \deg D$ and $\sup f_*(D) = f_*(\sup(D))$. If k is algebraically closed, so that divisors correspond to points, then $f_*([x])$ is just $[f(x)]$. Now assume that f is a finite map. Then we have a map $f^* : \operatorname{Div}(Y) \to \operatorname{Div}(X)$, (almost) characterized by two properties: $\deg(f^*(D)) = \deg(f)\deg(D)$ and $\sup(f^*(D)) = f^{-1}(\sup(D))$. (Here f^{-1} is just taken at the topological level.) If k is algebraically closed, then for $y \in Y(k)$ we have

$$f^*([y]) = \sum_{x \in X(k)} \operatorname{len}_x(X_y)[x]$$

where here $X_y = X \times_Y y$. Note that X_y is a finite subscheme of X, but may not be reduced.

Now let $f = (f_1, f_2)$ be a correspondence of X with total space C. We define a map $f_* : \operatorname{Div}(X) \to \operatorname{Div}(X)$ by $(f_2)_* f_1^*$. (If C is not regular, use $\widetilde{f}$.) This map carries principal divisors into principal divisors and therefore induces a map $f_* : \operatorname{Pic}(X) \to \operatorname{Pic}(X)$, and a map $f_* : \operatorname{Jac}(X) \to \operatorname{Jac}(X)$. (We have only defined these maps on field points, but they exist as maps of schemes.) Let $g : X \to X$ be a finite map of curves. The $f = (\mathrm{id}, g)$ and $f' = (g, \mathrm{id})$ are correspondences of X. The map f_* of $\operatorname{Jac}(X)$ coincides with the map g_*, while the map $(f')_*$ is the dual to g_*.

3.5. **The Eichler-Shimura identity.** There are two natural maps $p_1, p_2 : X_0(Np) \to X_0(N)$, taking (E, G) to E or E/G, where here G is a subgroup of order p and the level N structure is implicit. These two maps define a correspondence from $X_0(N)$ to itself over $\mathbf{Z}[1/N]$, called the *Hecke correspondence*, and denoted T_p. The Eichler–Shimura identity computes this correspondence in characteristic p:

Proposition 9. *We have $T_p = (F, 1) + (1, F)$ as correspondences on $X_0(N)$. Thus $(T_p)_* = F + V$ as endomorphisms of $J_0(N)$. (All of this takes place over $\mathbf{F}_p$.)*

Proof. Recall that we have defined maps $i_1, i_2 : X_0(N) \to X_0(pN)$ and that the map

$$i_1 \amalg i_2 : X_0(N) \amalg X_0(N) \to X_0(pN)$$

is the normalization of $X_0(pN)$. Thus, since we are allowed to replace the total space of a correspondence with its normalization, the correspondence T_p is just the sum of the correspondences $(p_1 \circ i_1, p_2 \circ i_1)$ and $(p_1 \circ i_2, p_2 \circ i_2)$, each with total space $X_0(N)$. Some short computations show that

$$p_1 \circ i_1 = F, \qquad p_1 \circ i_2 = \mathrm{id}, \qquad p_2 \circ i_1 = \mathrm{id}, \qquad p_2 \circ i_2 = F.$$

Thus $T_p = (F,1) + (1,F)$, as claimed. The correspondence $(1,F)$ induces the map F on the Jacobian while the correspondence $(F,1)$ induces the dual map V. This completes the proof. □

4. Applications to modular Galois representations

We have defined correspondences T_p on $X_0(N)$ for all primes p not dividing N. These correspondences induce endomorphisms of the abelian variety $J_0(N)$ and generate a commutative subalgebra $\mathbf{T}$ of $\mathrm{End}(J_0(N))$ which is finite over $\mathbf{Z}$. Let f be a cuspidal weight 2 newform of level N. Let $K_f \subset \mathbf{C}$ be the coefficient field of f. The action of $\mathbf{T}$ on f determines a homomorphism $\lambda : \mathbf{T} \to K_f$. Let $\mathfrak{a}$ be the kernel of λ, an ideal of $\mathbf{T}$, and let A_f be the quotient of $J_0(N)$ by $\mathfrak{a}J_0(N)$. We wish to understand A_f as best we can. We begin by computing its dimension:

Proposition 10. *The dimension of A_f is the degree of K_f.*

Proof. The dimension of A_f is the dimension of the space of global 1-forms on it. Global 1-forms on A_f correspond to global 1-forms on $J_0(N)$ which are killed by $\mathfrak{a}$. The latter space is precisely the space of modular forms with eigenvalue some $\mathrm{Aut}(\mathbf{C})$ conjugate of λ. Thus the dimension of this space is the number of conjugates of λ, which is equal to the degree of K_f. □

The abelian variety A_f actually has an action of K_f by isogenies, that is, there is a canonical map $K_f \to \mathrm{End}(A_f) \otimes \mathbf{Q}$. We can therefore regard the Tate module $T_\ell(A_f)$ as a two dimensional vector space over $K_f \otimes \mathbf{Q}_\ell$. The following proposition (combined with Faltings' theorem on isogenies of abelian varieties) determines A_f up to isogeny.

Proposition 11. *Let p be a prime not dividing N. Then A_f has good reduction at p and the trace of Frob_p on any Tate module $T_\ell A_f$ with $\ell \neq p$ is equal to a_p, the eigenvalue of T_p on f.*

Proof. We consider only the case $K = \mathbf{Q}$ for simplicity. The abelian variety $J_0(N)$ has good reduction at p since $X_0(N)$ is a smooth proper curve over $\mathbf{Z}[1/N]$, from which it follows that the quotient A_f does as well. Working modulo p, we have $T_p = F + V$ on $J_0(N)$, which implies the same on the quotient A_f. On this quotient, T_p acts by multiplication by the integer $a_p = \lambda(T_p)$. We thus find that $F^2 - a_p F + p = 0$ on A_f, which shows that a_p is the trace of F on $T_\ell A_f$. (There is a slight gap in this proof. In this section, we have simply stated that $\mathbf{T}$ acts on $J_0(N)$ without explaining how. To prove the Eichler–Shimura correspondence, we used a precise definition of the action of T_p on $J_0(N)$, and equated it with the action of the correspondence gotten by passing to the normalization of the total space. One must show that these two actions of T_p on $J_0(N)$ coincide to have a complete proof. We leave these details to the reader.) □

We now turn to ordinarity. We say that f is *ordinary* at p if its T_p eigenvalue is a p-adic unit. We say that a Galois representation $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Q}_p)$ is *ordinary* at p if on inertia it is an extension of 1 by χ_p. Furthermore, we say that ρ is *ordinary crystalline* if this extension class is represented by a unit in Kummer theory. We now have the following result:

Proposition 12. *If f is ordinary at $p \nmid N$ then $T_p A_f$ is ordinary crystalline.*

Proof. We again assume $K = \mathbf{Q}$. As in the previous proposition, A_f has good reduction at p and $F^2 - a_p F + p$ holds as an endomorphism of $\overline{A}_f$, the reduction of A_f modulo p. Assume for the moment that $\overline{A}_f$ were supersingular. Then we would have $F^2 = 0$ on $\overline{A}_f[p]$ (since $F = V$ in the supersingular case and $FV = p$), and so $a_p F$ would be zero on $\overline{A}_f[p]$. Since a_p is a p-adic unit, and integer, this would imply that F vanishes on $\overline{A}_f[p]$. This is impossible since the kernel of F has order p. Thus $\overline{A}_f$ is ordinary. The result now follows from the following proposition. □

Proposition 13. *Let $E/\mathbf{Z}_p$ be an elliptic curve with ordinary reduction. Then $T_p E$ is crystalline ordinary.*

Proof. It is not difficult to see that $E[p]$ is an extension of $\mathbf{Z}/p$ by μ_p over $\mathbf{Z}_p^{\mathrm{un}}$. In fact, we have an extension

$$0 \to \mu_{p^n} \to E[p^n] \to \mathbf{Z}/p^n \to 0$$

over $\mathbf{Z}_p^{\mathrm{un}}$. This shows that $T_p E$ is an extension of 1 by χ_p, and thus ordinary. To see that it is crystalline, note that the above extension, regarded simply as an extension of sheaves of groups on the fppf site of $\mathbf{Z}_p^{\mathrm{un}}$,

defines an element of $H^1_f(\mathbf{Z}_p^{\mathrm{un}}, \mu_{p^n})$. Here H_f is cohomology in the fppf site. We can compute this group via Kummer theory. Since $H^1_f(\mathbf{Z}_p^{\mathrm{un}}, \mathbf{G}_m) = 0$, we have

$$H^1_f(\mathbf{Z}_p^{\mathrm{un}}, \mu_{p^n}) = (\mathbf{Z}_p^{\mathrm{un}})^\times / ((\mathbf{Z}_p^{\mathrm{un}})^\times)^{p^n}.$$

This isomorphism is compatible with Kummer theory over $\mathbf{Q}_p^{\mathrm{un}}$, which shows that the extension class for T_pE is represented by a p-adic unit. □

5. Galois representations coming from Hilbert modular forms

Let f be a Hilbert cuspidal eigenform of parallel weight 2 for a totally real field F. We know that one can associate a Galois representation to f, and that its local properties are determined by those of f. Can this be proved in the same manner as the classical modular case?

If F has odd degree over $\mathbf{Q}$ or f is square-integrable at some finite place, then the Jacquet-Langlands correspondence shows that f can be transferred to a Shimura curve X. The curve part is the key point. One can then construct a quotient of the Jacobian of X, as we did for the Jacobian of $J_0(N)$, and obtain an abelian variety A_f. Everything goes through as before. (Some points may even be more simple, as Shimura curves are naturally compact — that is, there is no need to add cusps.)

When F has even degree over $\mathbf{Q}$ and f is principal series at all finite places, this procedure does not work. In fact, it is not known if the Galois representation associated to f appears as the Tate modular of an abelian variety, though I imagine this is expected. However, the Galois representation has been constructed and its local properties established, by more indirect means. If f is ordinary, then it can be put into a p-adic family. Other members of this family have Galois representations which are easier to construct, and the representation for f can be constructed as a limit. If f is not ordinary, this approach is not feasible. However, one can find enough congruences between f and forms for which the Galois representation is known to exist to construct the Galois representation of f. This was Richard Taylor's thesis.

References

[KM] N. Katz and B. Mazur, *Arithmetic moduli of elliptic curves*.
[DR] P. Deligne and M. Rapoport, *Les schémas de modules de courbes elliptiques*.

Existence of Taylor-Wiles Primes

Michael Lipnowski

Introduction

Let F be a totally real number field, $\overline{\rho} = \overline{\rho_f} : G_F \to GL_2(k)$ be an odd residually modular representation (odd meaning that complex conjugation acts as $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ for every archimedean place).
Let St be the set of places where ρ_f is Steinberg, S_p is the set of places over p, S_∞ the set of archimedean places of F, , and assume it is unramified everywhere else. For the purposes of this write up, all that matters is that $St \cup S_p$ is a finite set of finite places.
Our is to construct certain auxillary sets of places Q of F which have associated deformation rings R_Q. Q will consist of so called *Taylor-Wiles Places.*

Definition. *A place v of F is a* ***Taylor-Wiles place*** *if it satisfies the following conditions.*

- $v \notin S \cup S_p$.
- $Nv \equiv 1\ (p)$.
- *The eigenvalues of $\overline{\rho}(Frob_v)$ are distinct and belong to k.*

Let $R^{\square,\chi}_{Q \cup St \cup S_p}$ be the universal framed deformation ring unramified outside of $Q \cup St \cup S_p$ with fixed determinant $\chi = \chi_p$, the p-adic cyclotomic character.
Let $L^{\square}$ be the completed tensor product of the universal framed local deformation rings at $v \in St \cup S_p$ of fixed determinant ψ_v and $B^{\square}$ the completed universal product of their Steinberg quotients (for $v \in St$,) and their ordinary-crystalline quotients for $v \in S_p$.
Let $R^{\square}_Q = R^{\square,\chi}_{Q \cup St \cup S_p} \otimes_{L^{\square}} B^{\square}$. This represents the universal framed deformation $\rho : G_F \to GL_2(R_Q)$ of $\overline{\rho}$ unramified outisde of $Q \cup St \cup S_p$ which is Steinberg at St and ordinary-crystalline at S_p.

Although we do allow ramification at Q, the Taylor-Wiles conditions control it tightly.
Let v be a Taylor-Wiles place and consider $\rho|_{G_{F_v}}$.
$\overline{\rho}$ is unramified at v. So, $\rho(I_v)$ lands inside the 1-units of $GL_2(R_Q)$, which is a pro-p group. But the wild inertia group $W_v \subset I_v$ is a pro-v group and so it gets killed. Thus, the reduction is tamely ramified at v. Even better,

Lemma. *$\rho|_{G_{F_v}}$ is a sum of two (tamely ramified) characters $\eta_1 \oplus \eta_2$.*

Proof. The tame galois group is generated by $\sigma = Frob_v$ and the group I_v. For every $\tau \in I_v$, we have the relation

$$\sigma\tau\sigma^{-1} = \tau^{Nv}. \quad (*)$$

By the Taylor-Wiles assumption on Frobenii, $\overline{\rho}(\sigma)$ has distinct eigenvalues. By Hensel's lemma, we may lift $\overline{\rho}(\sigma)$ so that $\rho(\sigma)$ is diagonal, say $\begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$,
with respect to some (possibly different) basis. With respect to this basis, express

$$\rho(\tau) = 1 + \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

For some $a, b, c, d \in m_Q$. Apply ρ to $(*)$ and expand to get

$$1 + \begin{pmatrix} a & b\alpha\beta^{-1} \\ c\beta\alpha^{-1} & d \end{pmatrix} = \sum_{k=0}^{Nv} \binom{Nv}{k} \begin{pmatrix} a & b \\ c & d \end{pmatrix}^k.$$

Note that for $k \geq 2$, the top right and bottom left entries of the right side summands lie in $m_Q(b, c)$. Thus comparing with these entries on the left side,

$$b(\alpha\beta^{-1} - Nv), c(\beta\alpha^{-1} - Nv) \in m_Q(b, c).$$

But α and β are residually distinct, by assumption. Then by the congruence property of TW places

$$\alpha\beta^{-1} - Nv, \beta\alpha^{-1} - Nv \neq 0 \ (p)$$

implying that both terms are units in R_Q. Thus, $(b, c) \subset m_Q(b, c)$. By Nakayama's Lemma, this implies that $b = c = 0$. Since τ was aribitrary, the claim follows. □

$\mathcal{O}[\Delta_Q]$ Structure on $R_Q^{\square}$

We have just shown that $\rho|_{G_{F_v}}$ is a sum of two (tamely ramified) characters $\eta_1 \oplus \eta_2$. Choose one, say η.
We know that $\eta|_{I_v}$ has pro-p image. Also by class field theory, it determines a character $\eta' : O_v^{\times} \to R_Q^{\square\times}$. As the 1-units are pro-v, this is really a map $\eta' : (O_v/v)^{\times} \to R_Q^{\square\times}$ which factors through the maximal p-power quotient of $(O_v/v)^{\times}$. Call this maximal p-power quotient Δ_v. Let $\Delta_Q = \prod_{v \in Q} \Delta_v$. Our choice of η defines an action of Δ_Q on R_Q, thus giving R_Q the structure of an $\mathcal{O}[\Delta_Q]$-module.

We still haven't constructed the set of primes Q. Actually, we want to construct a family of such $Q = Q_n$ of the following sort:
For fixed positive integers g, h satisfying $\dim B^{\square} = 1 + h + l - g$ (remember that $B^{\square}$ is the framed ring of Steinberg and ord-cryst conditions),

- $|Q_n| = h$
- $Nv = 1 \ (p^n)$
- $R_{Q_n}^{\square}$ is topological generated by g elements over $B^{\square}$.

Note that the congruence condition $Nv = 1\ (p^n)$ means that Δ_v is $p-$power cyclic of order divisible by p^n. Thus, after a choice of generators for these cyclic groups, the $\mathcal{O}[\Delta_Q]$-module structure on $R^{\square}_Q$ is equivalently an $\mathcal{O}[[T_1, ..., T_h]]/((T_1+1)^{p^{a_1}} - 1, ..., (T_h+1)^{p^{a_h}} - 1)$-module structure, where all $a_i \geq n$.
There are no obvious maps between the R_{Q_n}. But by the magic of the patching, we will find a subset of the R_{Q_n} which form a kind of inverse system with limit $R^{\square}_{\infty}$. We dream that by "letting $n \to \infty$", we'll give $R^{\square}_{\infty}$ the structure of a free $\mathcal{O}[[T_1, ..., T_h]]$-module.
A couple remarks about these conditions:

1) The explicit values

$$\begin{aligned} h &= \dim H^1(G_{F,St\cup S_p}, ad^0\overline{\rho}(1)) \\ g &= h - [F:\mathbb{Q}] + |St| + |S_p| - 1 \end{aligned}$$

will suffice.

2) Our stipulation that $\dim B^{\square} = 1 + h + l - g$ will only appear natural once we dive into the patching argument.

3) The g we will construct is actually the relative topological dimension of $R^{\square}_{Q_n}$ over $L^{\square}$, which will certainly suffice.

Construction of the TW Sets

From now on, we will assume that

$$\overline{\rho}|_{G_{F(\zeta_p)}} \text{is absolutely irreducible.}$$

This cheaply implies the following apparently much stronger fact.

Lemma. $\overline{\rho}|_{G_{F(\zeta_{p^n})}}$ *is absolutely irreducible.*

Proof. Our standing assumption is that $\overline{\rho}|_{G_{F(\zeta_p)}}$ is absolutely irreducible.
Note that $H = G_{F(\zeta_{p^n})}$ is a normal subgroup of $G = G_{F(\zeta_p)}$. Thus, the restriction $\overline{\rho}|_H$ is semisimple. Indeed, if W is an invariant subspace, then

$$\bigoplus_{G/H - 1.H} gW$$

is an invariant complement.
Suppose $\overline{\rho}|_H$ is not irreducible. Then it is the direct sum of two characters. Since V, as a G-module, is absolutely irreducible, G/H must permute these characters transitively. But G/H is a p-group, and so it cannot act transitively on a 2 element set (for any $p > 2$, which we have assumed). Thus, the two characters are the same.
This implies that every line in V is stabilized by H. But there are $|\mathbb{P}(V)(k)| = |k| + 1$ of them. So the number of them is prime to p. Hence, some orbit of G/H on the set of k-lines in V has size prime to p. But the size of the orbit must also divide $|G/H|$, which is p-power. Hence, this orbit has size 1, i.e. there is an H-stable line which is G/H-stable. This line is then G-stable, contradicting the irreducibility of V.
The same argument carries out mutatis mutandis after first making a finite extension of the ground field k of V. Thus, $\overline{\rho}|_H$ is indeed absolutely irreducible. □

We'll now prove our main lemma of interest.

Theorem (DDT, Lemma 2.49). *Let $h = \dim H^1(G_{F,St\cup S_p}, ad^0(\overline{\rho}(1)))$. For every n, we can construct a set Q_n of Taylor-Wiles places, i.e.*

(1) For each $v \in Q_n, Nv = 1\ (p^n)$.

(2) For each $v \in Q_n, \overline{\rho}(Frob_v)$ has distinct k-rational eigenvalues.

(3) $|Q_n| = h$.

Proof. An easy calculation shows that if $\overline{\rho}(Frob_v)$ is a Taylor-Wiles place, then $\dim H^1(k_v, ad^0(\overline{\rho})(1)) = 1$.
Indeed, for any σ in $G_{F,St\cup S_p}$, if σ has (generalized) eigenvalues α, β then $ad^0(\overline{\rho})(\sigma)$ has (generalized) eigenvalues $1, \alpha\beta^{-1}, \beta\alpha^{-1}$. Thus, if $\overline{\rho}(Frob_v)$ has distinct eigenvalues, the space $ad^0(V)/(Frob_v - 1)ad^0(V)$ is one dimensional. Since a v-unramified cocycle is uniquely determined by its value on $Frob_v$, we get that $\dim H^1(k_v, ad^0(\overline{\rho})(1)) = 1$.
Thus, it suffices to show that the restriction map

$$H^1(G_{F,St\cup S_p}, ad^0(\overline{\rho})(1)) \to \oplus_{v\in Q_n} H^1(k_v, ad^0(\overline{\rho})(1))$$

is an isomorphism. Then equating dimensions shows that condition (3) is fulfilled.
To do this, it suffices to show that for any global cocycle ψ there exists a $v = v_\psi$ satisfying (1) and (2) such that $res_v(\psi) \neq 0$. For then we could apply this to the elements of a basis (of size h) for the left side, and the corresponding set of places $\{v_\psi\}$ would consistute a TW set.
Instead we'll show that we can find $\sigma \in G_{F,St\cup S_p}$ satisfying the following:

(1') $\sigma|_{G_{F(\zeta_p)}} = 1$.

(2') $ad^0\overline{\rho}(\sigma)$ has an eigenvalue other than 1.

(3') $\psi(\sigma) \notin (\sigma - 1)ad^0\overline{\rho}(1)$.

Indeed, all three of the above conditions are open conditions in $G_{F,St\cup S_p}$. But by the Chebotarev density theorem, we any non-empty open set contains some $Frob_v$. This v will do.

Let F_0 be the fixed field of the kernel of $ad^0\overline{\rho}$ and let $F_m = F_0(\zeta_{p^m})$.

Claim. *$\psi(G_{F_n})$ is non-zero.*

Later, we'll even show that its k-span is a non-zero $Gal(F_n/F(\zeta_{p^n}))$-submodule of $ad^0\overline{\rho}$. From this, we can leverage information from the irreducibility of $\overline{\rho}|_{G_{F(\zeta_{p^n})}}$ just proven.

Proof. In this claim and what follows, assume $n > 0$ so that the cyclotomic character is trivial when restricted to G_{F_n}. There is an inflation-restriction sequence

$$0 \to H^1(G_{F_n/F}, ad^0\overline{\rho}(1)) \xrightarrow{inf} H^1(G_F, ad^0\overline{\rho}(1)) \xrightarrow{res} H^1(G_{F_n}, ad^0\overline{\rho}(1)).$$

Thus, it suffices to prove that the leftmost term is zero. For then, $\psi|_{G_{F_n}}$ is a non-zero cohomology class, and so is certainly not identically 0.
We can sandwich the left most term in another inflation-restriction sequence:

$$0 \to H^1(G_{F_0/F}, (ad^0\overline{\rho}(1))^{G_{F_0}}) \xrightarrow{inf} H^1(G_{F_n/F}, ad^0\overline{\rho}(1)) \xrightarrow{res} H^1(G_{F_n/F_0}, ad^0\overline{\rho}(1))^{G_{F_0/F}}. \ (*)$$

where the action of $g \in G_{F_0/F}$ on the third term is given by $\eta \mapsto (h \mapsto g^{-1}\eta(ghg^{-1}))$.

- Third term of (*)
  There is a restriction-corestriction sequence

$$H^1(G_{F_n/F_0}, ad^0\overline{\rho}(1)) \xrightarrow{res} H^1(G_{F_n/F_1}, ad^0\overline{\rho}(1)) \xrightarrow{cores} H^1(G_{F_n/F_0}, ad^0\overline{\rho}(1))$$

  and the composition is multiplication by $|G_{F_1/F_0}|$. This number is $\leq p-1$ and so is prime to p. Hence, res is injective. It also sends $G_{F_0/F}$-invariants to $G_{F_0/F}$-invariants. Thus, it suffices to show that $H^1(G_{F_n/F_1}, ad^0\overline{\rho}(1))^{G_{F_0/F}}$ is zero.

  - G_{F_n/F_1} is naturally a subgroup of the commutative quotient $G_{F(\zeta_{p^n})/F}$ of G_F (given just by restricting automorphisms to $F(\zeta_{p^n})$). The conjugation action is compatible with this restriction. Thus the conjugation action on G_{F_n/F_1} is trivial since the latter quotient of G_F is abelian.
    Note that G_{F_n/F_1} acts trivially on $ad^0\overline{\rho}(1)$. Hence,

$$H^1(G_{F_n/F_1}, ad^0\overline{\rho}(1))^{G_{F_0/F}} = Hom(G_{F_n/F_1}, ad^0\overline{\rho}(1))^{G_{F_0/F}} = Hom(G_{F_n/F_1}, ad^0\overline{\rho}(1)^{G_{F_0/F}}).$$

    But $ad^0\overline{\rho}(1)^{G_{F_0/F}} = 0$.
    Indeed, any $G_{F(\zeta_{p^n})}$-invariant element of $ad^0\overline{\rho}(1)$ is equivalently a trace 0 intertwining operator $V \to V(1)$ (V the underlying vector space of ad^0). But $n > 0$, so the action of the cyclotomic character is trivial. So this is actually an intertwining operator $V \to V$. But V is an irreducible $G_{F(\zeta_{p^n})}$-module, and so any self-intertwining operator is scalar and so must be 0 by our trace 0 assumption ($p > 3$ by our standing assumptions).

  Hence, the third term of $(*)$ is 0.

- First term of (*)

  - $(ad^0\overline{\rho}(1))^{G_{F_0/F}}$ is trivial unless $F_0 \supset F(\zeta_p)$. This is because for any place v with $Nv \neq 1\ (p), ad^0\overline{\rho}(Frob_v)$ fixes something but $\chi_p(Frob_v) \neq 1$. So, we assume that

$$G_{F_0/F} \to G_{F(\zeta_p)/F} \to 0.$$

    In particular, $G_{F_0/F}$ has a non-trivial quotient and so is not a non-abelian simple group.
  - Since $(ad^0\overline{\rho}(1))^{G_{F_0/F}}$ has p-power order, we also have an injection

$$0 \to H^1(G_{F_0/F}, (ad^0\overline{\rho}(1))^{G_{F_0}}) \xrightarrow{res} H^1(P, (ad^0\overline{\rho}(1))^{G_{F_0}}),$$

    where P is the Sylow p-subgroup of $G_{F_0/F}$. Thus, we can assume that P is non-trivial, i.e. that p divides $|G_{F_0/F}|$.
  - Finally, since F_0 is the field cut out by $ad^0\overline{\rho}, G_{F_0/F}$ is isomorphic to the projective image of $\overline{\rho}$.

We can put these facts to good use in conjunction with an explicit characterization of finite subgroups of $PGL_2(\overline{\mathbb{F}}_p)$.

List of Finite Subgroups H of $PGL_2(\overline{\mathbb{F}}_p)$ [**EG, II.8.27**]

– H is conjugate to a subgroup of the upper triangular matrices.
– H is conjugate to $PGL_2(\mathbb{F}_{p^r})$ or $PSL_2(\mathbb{F}_{p^r})$ for some $r \geq 1$.
– H is isomorphic to A_4, A_5, S_4, or $D_{2r}, p \nmid r$ for $r \geq 2$. Furthermore, if H is isomorphic to $D_{2r} = \langle s, t | s^2 = t^r = 1, sts = t^{-1} \rangle$, then it is conjugate to the image of

$$s \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} t \mapsto \begin{pmatrix} \zeta & 0 \\ 0 & 1 \end{pmatrix},$$

where ζ is a primitive r^{th} root of unity.

We can eliminate all of these possibilities, one by one.

– The projective image H cannot be conjugate to a subgroup of the upper triangular matrices, for then $\overline{\rho}|_{G_{F(\zeta_p)}}$ would not be absolutely irreducible.
– Our assumptions $p > 5$ and p divides $|G_{F_0/F}|$ preclude the possibilities $H \cong A_4, A_5, S_4, D_{2r}, p \nmid r$.
– $PSL_2(\mathbb{F}_{p^r})$ is simple for $p > 5$. Thus, it cannot have a quotient, namely $G_{F(\zeta_p)/F}$, which is non-trivial.
– Suppose $H = im(\overline{\rho}) \cong PGL_2(\mathbb{F}_{p^r})$. The only non-trivial quotient of $PGL_2(\mathbb{F}_{p^r})$ is order 2. But $G_{F_0/F}$ cannot have a quotient of order 2.
If it did, there would be an exact sequence

$$1 \to Z \to im(\overline{\rho}) \to im(ad^0(\overline{\rho})) \to 1,$$

with Z a central subgroup of $GL_2(k)$ and $im(ad^0(\overline{\rho}))$ either order 1 or 2. But then any pre-image A of the non-trivial element of $im(ad^0(\overline{\rho}))$ and Z generate $im(\overline{\rho})$. But A has an invariant subspace (possibly after a quadratic extension). So that means $im(\overline{\rho})$ does too, contradicting the absolute irreducibility of $\overline{\rho}$.

Since none of these are possible, we must have the first term of $(*)$ being 0 after all.

We conclude that the second term of $(*)$ is 0 as well, which is what we wanted; this proves that $\psi(G_{F_n})$ is indeed non-zero. □

We can say more. For $\tau, \tau' \in G_{F_n}, \sigma \in G_{F(\zeta_{p^n})}$, repeated use of the cocycle relation gives

$$\begin{aligned} \psi(\sigma\tau\sigma^{-1}) &= \psi(\sigma) + \psi(\tau\sigma^{-1}) \\ &= \psi(\sigma) + \sigma\psi(\tau) + \sigma\tau\psi(\sigma^{-1}) \\ &= \psi(\sigma) + \sigma\psi(\tau) + \sigma\psi(\sigma^{-1}) = \sigma\psi(\tau). \end{aligned}$$

Note: the second last equality holds because τ acts trivially on $ad^0\overline{\rho}(G_{F_n})$. Also,

$$\psi(\tau) + \psi(\tau') = \tau'\psi(\tau) + \psi(\tau') = \psi(\tau\tau').$$

Thus, the k-span of $\psi(G_{F_n})$ is in fact a non-zero $G_{F_n/F(\zeta_{p^n})}$-submodule of $ad^0\overline{\rho}$.

Next, we'll find an element $g \in G_{F_n/F(\zeta_{p^n})}$ such that $\overline{\rho}(g)$ has distinct eigenvalues and which fixes an element of $k.\psi(G_{F_n})$. We do this by the explicit classification of possible projective images, i.e. we'll show that for any subgroup H which could possibly be the projective image of $\overline{\rho}$, there is an element of H with distinct eigenvalues which fixes an element of $k.\psi(G_{F_n})$.

- Note first that if we can prove the result for some subgroup $H \subset H'$, then it true for putative projective image H' as well. Also, the "exceptional" cases A_4, S_4, and A_5 all contain D_4 and the projective image cannot be contained in an upper triangular subgroup (due to the absolute irreducibility of $\overline{\rho}|_{G_{F(\zeta_{p^n})}}$. Thus, in view of the preceding classification of finite subgroups of $\mathrm{PGL}_2(\overline{\mathbb{F}}_p)$, it suffices to check the following cases:

- $\underline{PSL_2(\mathbb{F}_{p^r})}$
  ad^0 is simple under the action of $PSL_2(\mathbb{F}_{p^r})$. Thus, $k.\psi(G_{F_n}) = ad^0$ and
  $\begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix}$ fixes $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \in ad^0 = k.\psi(G_{F_n})$. Since $p > 5$, we can certainly find $\alpha \neq \alpha^{-1}$.

- $\underline{D_4}$
  ad^0 decomposes as $V_1 \oplus V_2 \oplus V_3$, where

$$V_1 = \left\langle \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\rangle, V_2 = \left\langle \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle, V_3 = \left\langle \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle.$$

  D_4 acts as ± 1 on each V_i. Furthermore, by our explicit description of the image of dihedral groups, each non-trivial element has distinct eigenvalues (of ± 1). Since the only possible invariant subspaces of ad^0 are then $\oplus_{i \in I} V_i$ for some $I \subset \{1, 2, 3\}$, it follows that some element $h \in D_4$ with distinct eigenvalues fixes an element of $k.\psi(G_{F_n})$.

- $\underline{D_{2r}, r \text{ odd}}$
  ad^0 decomposes as $W_1 \oplus W_2$ where

$$W_1 = \left\langle \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle, W_2 = \left\langle \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\rangle.$$

  W_1 is fixed by $\begin{pmatrix} 1 & 0 \\ 0 & \zeta \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ fixes $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

  Since $ad^0 = W_i$ or $W_1 \oplus W_2$, it follows again that some $h \in D_{2r}$ with distinct eigenvalues fixes an element of $k.\psi(G_{F_n})$.

Having found such a g, it must certainly fix a non-zero element of $\psi(G_{F_n})$ itself, say $\psi(\tau_0)$.

- Indeed, as an $\mathbb{F}_p$-vector space, the $k.\psi(G_{F_n})$ is isomorphic to $k \otimes_{\mathbb{F}_p} \psi(G_{F_n})$. But then if $k_1, ..., k_m$ forms a basis for k over $\mathbb{F}_p$, we can express the fixed element m of $k.\psi(G_{F_n})$ as $m = k_1\psi(\tau_1) + ... + k_n\psi(\tau_n)$, where at least one $\psi(\tau_i) \neq 0$. If m is fixed by g, then

$$k_1((g-1)\psi(\tau_1)) + ... + k_n((g-1)\psi(\tau_n)) = 0.$$

  But linear independence implies that $(g-1)\psi(\tau_i) = 0$, which is what we wanted.

Choose a lift σ_0 of g to the absolute Galois group.
For $\tau \in G_{F_n}$, we have

$$\psi(\tau\sigma_0) = \tau\psi(\sigma_0) + \psi(\tau) = \psi(\sigma_0) + \psi(\tau).$$

- If $\psi(\sigma_0) \notin (\sigma_0 - 1)(ad^0\overline{\rho}(1))$, then take $\tau = 1$.
- Otherwise, choose $\tau = \tau_0$. For this choice, $\psi(\tau_0) \notin (\sigma_0 - 1)ad^0\overline{\rho}(1)$. For suppose $(\sigma_0 - 1)x = \psi(\tau_0) \neq 0$. Applying $\sigma_0 - 1$ to both sides, our construction of τ_0 gives

$$(\sigma_0 - 1)^2 x = (\sigma_0 - 1)\psi(\tau_0) = 0.$$

But σ_0 acting on ad^0 is semisimple and has eigenvalue 1 with multiplicity 1 (since $\overline{\rho}(\sigma_0)$ has distinct eigenvalues) . Thus, $(\sigma_0 - 1)x = 0$, implying that $\psi(\tau_0) = 0$, contrary to our construction.

Thus, in both cases

$$\psi(\tau\sigma_0) \notin (\sigma_0 - 1)ad^0\overline{\rho}(1) = (\tau\sigma_0 - 1)ad^0\overline{\rho}(1).$$

So we've finally constructed the element $\sigma = \tau\sigma_0$ that we sought in the first place. □

Number of Topological Generators for $R^{\square,\chi}_{Q_n \cup St \cup S_p}$ over $L^{\square}$

We now have all of the pieces in place to compute the relative tangent space dimension of $R^{\square,\chi}_{Q_n \cup St \cup S_p}$ over $L^{\square}$, both defined as in the introduction.

Lemma (FFGS, 3.2.2). *Let $h^1(G_{F,St\cup S_p\cup S_\infty}, ad^0(V))$ denote the k-dimension of*

$$\ker(H^1(G_{F,St\cup S_p\cup S_\infty}, ad^0(V)) \to \prod_{v\in St\cup S_p} H^1(G_{F_v}, ad^0(V))).$$

For $v \in St \cup S_p$, let $\delta_v = \dim_k H^0(G_{F,St\cup S_p\cup S_\infty}, adV)$ and $\delta_F = \dim_k H^0(G_{F,St\cup S_p\cup S_\infty}, adV)$. Then $R^{\square,\chi}_{F,St\cup S_p\cup S_\infty}$ is the quotient of a power series ring over $L^{\square}$ in

$$g = h^1(G_{F,St\cup S_p\cup S_\infty}, ad^0(V)) + \sum_{v\in St\cup S_p} \delta_v - \delta_F.$$

variables.

Proof. Let our vector space V have fixed basis β.
An element of the relative tangent space corresponds to a deformation of V to a finite free $k[\epsilon]$-module $\tilde{V}$ together with a choice of bases $\tilde{\beta}_v$ lifting β such that for each $v \in St \cup S_p$, the pair $(\tilde{V}|_{G_{F_v}}, \beta_v)$ is isomorphic to $(V \otimes_k k[\epsilon], \beta \otimes_k 1)$.
For fixed choices of bases, the space of such deformations is given by

$$ker(H^1(G_{F,St\cup S_p\cup S_\infty}, ad^0(V)) \to \prod_{v\in St\cup S_p} H^1(G_{F_v}, ad^0(V))).$$

Given such a deformation, $\tilde{V}$, the space of possible choices for a bases is the space of G_{F_v} automorphisms of $(V \otimes_k k[\epsilon], \beta \otimes_k 1)$; such an automorphism reduces to 1 mod (ϵ) and so is of the form $1 + \epsilon M$ for some G_{F_v}-equivariant $M \in ad(V)$, i.e. $M \in H^0(G_{F_v}, adV)$.
The same reasoning shows that two collections $\{\beta_v\}_{v\in St\cup S_p}$ and $\{\beta'_v\}_{v\in St\cup S_p}$ determine the same framed deformation if they differ by an element of $H^0(G_{F,St\cup S_p\cup S_\infty}, adV)$. The lemma follows.

□

Now we compute this h^1, the dimension of a Selmer group, via the Wiles Product formula.

Lemma (FFGS, 3.2.5). *Set $g = \dim_k H^1(G_{F,S_p\cup St}, ad^0\overline{\rho}(1)), ad^0\overline{\rho}(1)) - [F:\mathbb{Q}] + |St| + |S_p| - 1$. For each positive integer n, there is a finite set of primes Q_n of F which is disjoint from $St \cup S_p$ and such that*

(1) If $v \in Q_n$, then $Nv = 1\ (p^n)$ and $\overline{\rho}(Frob_v)$ has distinct eigenvalues.

(2) $|Q_n| = \dim_k H^1(G_{F,S_p\cup St}, ad^0\overline{\rho}(1))$. Also, $R^{\square}_{Q_n}$ is topolgoically generated by g elements as a $B^{\square}$-algebra.

Proof. We define a set of local conditions to compute this relative dimension, the dimension of a Selmer group. Namely, let

$$H^1_{\mathcal{L}_v} = \begin{cases} 0 & \text{if } v \in St \cup S_p \\ H^1(G_{F_v}, ad^0\overline{\rho}) & \text{otherwise.} \end{cases}$$

Write $H^1_{\mathcal{L}_{Q_n}}$ (resp. $H^1_{\mathcal{L}^{\perp}_{Q_n}}$) for the set of classes which restrict to $H^1_{\mathcal{L}_v}$ (resp. $H^1_{\mathcal{L}^{\perp}_v}$) for each $v \in St \cup S_p \cup Q_n$. ("$\perp$" denoting the annihilator under Tate local duality).
The main result from the previous section shows that we can find a set of primes Q_n satisfying condition (1) and the first part of condition (2). Furthermore, any class in $H^1_{\mathcal{L}^{\perp}_{Q_n}}$ restricts to 0 in $H^1(G_{F_v}, ad^0\overline{\rho}(1))$. By our choice of primes, this implies that $H^1_{\mathcal{L}^{\perp}_{Q_n}} = 0$.

By the Wiles Product Formula, we get

$$|H^1_{\mathcal{L}_{Q_n}}| = \frac{H^0(G_{F,St\cup S_p\cup S_\infty}, ad^0\overline{\rho})}{H^0(G_{F,St\cup S_p\cup S_\infty}, ad^0\overline{\rho}(1))} \prod_{v\in St\cup S_p\cup S_\infty} \frac{H^1_{\mathcal{L}_v}}{H^0(G_{F_v}, ad^0\overline{\rho})}.$$

- Global terms
  An element of $H^0(G_{F,St\cup S_p\cup S_\infty}, ad^0\overline{\rho})$ corresponds to a trace 0 self-intertwining operator of V. Since $\overline{\rho}|_{G_{F(\zeta_p)}}$ is absolutely irreducible, any self-intertwining operators are scalars. But the only trace 0 scalar matrix is 0 (for $p > 2$).
  Similarly, an element of $H^0(G_{F,St\cup S_p\cup S_\infty}, ad^0\overline{\rho}(1))$ corresponds to an intertwining operator $V \to V(1)$ between irreducible $G_{F(\zeta_p)}$-modules. Either they are not isomorphic, in which case only the 0 operator can intertwine them, or they are isomorphic, in which case the above paragraph applies.

- $v \in St \cup S_p$
  $ad^0(V)$ is a summand of $ad(V)$ for $p > 2$. So, the terms in the product corresponding to $v \in St \cup S_p$ in the product formula contribute $|k|^{1-\delta_v}$.

- $v \in S_\infty$

- $v \in Q_n$

  $$\frac{H^1(G_{F_v}, ad^0\overline{\rho})}{H^0(G_{F_v}, ad^0\overline{\rho})} = H^2((G_{F_v}, ad^0\overline{\rho})) \times \text{local Euler characteristic}^{-1}.$$

  The H^2 term equals $H^0(G_{F_v}, ad^0\overline{\rho}(1))$ by Tate local duality. The local Euler characteristic, which equals $[O_v : |ad^0(V)|O_v]^{-1}$ by the local Euler characteristic formula, is 1 since $|ad^0(V)|$ is prime to $v \in Q_n$. Hence, the product formula terms for $v \in Q_n$ equal $H^0(G_{F_v}, ad^0\overline{\rho}(1))$.

Since $\overline{\rho}(Frob_v)$ had distinct eigenvalues, there is a 1-dimensional subspace of $ad^0(V)$ fixed by $ad^0\overline{\rho}(Frob_v)$. Since $\overline{\rho}|_{G_{F_v}}$ is unramified, $H^0(G_{F_v}, ad^0\overline{\rho}(1))$ is 1-dimensional.

- $\underline{S_\infty}$

  By one of our standing assumptions, $\overline{\rho}$ is odd, i.e. for archimedean places $v, \overline{\rho(c)}$ can represented as a matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ with respect to some basis. Hence, $ad^0\overline{\rho}(c)$ is can be diagonalized to $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. But G_{F_v} is cyclic of order 2, generated by c. Hence, the space of cocycles is just $\ker(ad^0\overline{\rho}(c)+1)$, which is 2-dimensional, and the space of coboundaries is $im(ad^0\overline{\rho}(c)-1)$, which is 2-dimensional. Hence $H^1(G_{F_v}, ad^0\overline{\rho}) = 0$. Also, $H^0(G_{F_v}, ad^0\overline{\rho})$ is the 1-eigenspace of $ad^0\overline{\rho}(c)$, and so is 1-dimensional.

Adding everything together, we get

$$\begin{aligned} h^1(G_{F,St\cup S_p\cup S_\infty}, ad^0(V)) &= \dim_k H^1_{\mathcal{L}_{Q_n}} \\ &= 0 + \sum_{v\in St\cup S_p}(1-\delta_v) + \sum_{v\in Q_n} 1 + \sum_{v\in S_\infty} -1 \\ &= |St| + |S_p| - \sum_{v\in St\cup S_p}\delta_v + |Q_n| + [F:\mathbb{Q}] \\ &= |St| + |S_p| - \sum_{v\in St\cup S_p}\delta_v + \dim_k H^1(G_{F,St\cup S_p}, ad^0\overline{\rho}(1)) + [F:\mathbb{Q}] \end{aligned}$$

Combining with the previous lemma gives that

$$g = \dim_k H^1(G_{F,St\cup S_p}, ad^0\overline{\rho}(1)) + |St| + |S_p| + [F:\mathbb{Q}] - 1,$$

as desired. □

We can conclude that $R^{\square}_Q$ is generated by g elements as a $B^{\square}$ algebra as well. Thus, we are finally done our construction of TW primes.

References

DDT H. Darmon, F. Diamon, R. Taylor. *Fermat's Last Theorem.* Current Developments in Mathematics 1 (1995), International Press, pp. 1-157.

FFGS M. Kisin. *Moduli of Finite Flat Group Schemes and Modularity.* Annals of Math. 170(3) (2009), 1085-1180.

EG B. Huppert. *Endliche Gruppen I.* Grundlehren Math. Wiss. 134 (1983), Springer-Verlag, New York, Berlin, Heidelberg.

S. Shah. *Framed Deformation and Modularity.* Harvard Undergraduate Thesis (2009). Available at http://math.harvard.edu/theses/senior/shah/shah.pdf

Construction and properties of the modules for patching (Modularity 5.20.10)

Sam Lichtenstein

July 1, 2010

1 Introduction/Motivation

Recall that our ultimate goal is to prove a modularity lifting theorem (stated in Andrew's talk "Overview of the Taylor-Wiles method"), which we have reduced to showing an $R = \mathbf{T}$ theorem. That is, we have a surjection from a deformation ring R' to a Hecke algebra $\mathbf{T}_{\mathfrak{m}}$, and we want to prove it is an isomorphism.

Brandon will prove this next time. The idea of the proof – the patching argument – is to use framed versions $R^{\square}_{\mathfrak{n}}$ of the rings $R_{Q_{\mathfrak{n}}}$. Here the $Q_{\mathfrak{n}}$ form a Taylor-Wiles system; Mike established the existence of such last time. These each live over $R^{\square}_{\emptyset}$ (which is very close to the R' we care about):

$$R^{\square}_{\mathfrak{n}} \twoheadrightarrow R^{\square}_{\emptyset}.$$

On the $\mathbf{T}$ side we will have certain modified Hecke algebras $\mathbf{T}^{\square}_{\mathfrak{n}}$ which I will describe later. We'll construct maps $R^{\square}_{\mathfrak{n}} \twoheadrightarrow \mathbf{T}^{\square}_{\mathfrak{n}}$ lifting $R^{\square}_{\emptyset} \to \mathbf{T}^{\square}_{\emptyset}$, and certain $R^{\square}_{\mathfrak{n}}$-modules $M^{\square}_{\mathfrak{n}}$. By a pigeonhole principle sort of thing, we'll be able to pass to an inverse limit

$$R^{\square}_{\infty} \twoheadrightarrow \mathbf{T}^{\square}_{\infty}.$$

We'll have an inverse limit $M^{\square}_{\infty}$. We'll be able to show that $M^{\square}_{\infty}[\frac{1}{p}]$ is a faithful $R^{\square}_{\infty}[\frac{1}{p}]$-module, and some other nice things. Then we'll deduce that this faithfulness must have been true at level $\emptyset$, where the $R^{\square}_{\emptyset}[\frac{1}{p}]$-action was through (a framed version of) our map $R' \to \mathbf{T}_{\mathfrak{m}}$, which we can therefore conclude is injective, hence an isomorphism.

Phew! This argument is clearly a technological marvel on par with my iPhone. For such a thing to work, we need precise control over what happens on each level $\mathfrak{n}$ as we go up the tower. Specifically, Brandon will need to show that a certain collection of rings and moduled cooked up from the $R^{\square}_{\mathfrak{n}}$ and $M^{\square}_{\mathfrak{n}}$ form a "patching datum", meaning that they satisfy a collection of technical axioms that make the patching argument go through.

The goal of this talk is two-fold. One thing I need to do is define the relevant Hecke algebras $\mathbf{T}^{\square}_{\mathfrak{n}}$ and modules $M^{\square}_{\mathfrak{n}}$. To ruin the surprise, the module $M^{\square}_{\mathfrak{n}}$ will arise from a space of modular forms (which remember, are just functions on a finite set because we cleverly set things up that way) of suitable level, depending on the Taylow-Wiles set of primes $Q_{\mathfrak{n}}$. The other thing I need to do is show that the modules $M^{\square}_{\mathfrak{n}}$ satisfy nice properties, so that Brandon can show that this his patching data are actually patching data. So this is all really

an expansion of section 4 of the notes from Andrew's overview talk, if you're following along at home.

Remark 1.0.1. Although ultimately we'll use framed stuff for patching, mostly we'll deal with unframed stuff in this talk.

2 Setup and statment of what we're going prove

Throughout, $\mathcal{O}$ is the ring of integers of a $\mathfrak{p}$-adic field with residue field k (where $\overline{\rho}$ lives).

Recall that the quaternion algebra D is ramified exactly at the places St and all the archimedean places. Let's fix some notation. For a compact open subgroup $U \subset (D \otimes_F \mathbf{A}_F^f)^\times$ set

$$X(U) = D^\times \backslash (D \otimes_F \mathbf{A}_F^f)^\times / (U \cdot (\mathbf{A}_F^f)^\times).$$

Let

$$S(U) = \text{Functions}(X(U), \mathcal{O}).$$

This is the space of automorphic forms on $D^\times$ of weight 2 and level U. Let the bad primes for U be

$$\Sigma(U) = S_p \cup St \cup \{v|\infty\} \cup \{v : U_v \text{ is non-maximal}\},$$

and define the Hecke algebra $\mathbf{T}(U)$ to be the $\mathcal{O}$-subalgebra of $\text{End}(S(U))$ generated by the T_v for $v \notin \Sigma(U)$. (It comes from the double coset $U_v \left(\begin{smallmatrix} 1 & 0 \\ 0 & \varpi_v \end{smallmatrix}\right) U_v$.)

Now let's fix the "ground level" $U^\circ \subset (\mathbf{A}_F^f \otimes D)^\times$ [a compact open subgroup] for the construction. Picking up on a technical point Andrew mentioned, which will be relevant today, we need to choose a huge prime v_{aux} which has nothing to do with anything. In other words it should be outside of $St \cup S_p \cup \{v|\infty\} \cup \bigcup_{n \geq 1} Q_n$. We can certainly arrange this because, for example, all the primes in the Q_ns satisfy $Nv \equiv 1 \bmod p$. Now take U° to be the maximal compact for all places $v \neq v_{aux}$; we will specify $U^\circ_{v_{aux}}$ later.

Let $\mathbf{T} = \mathbf{T}(U^\circ)$

Recall that we have a modular lift ρ_f of our residual representation $\overline{\rho}$ which satisfies a bunch of nice properties. Via Jacquet-Langlands and our assumptions on ρ_f, the Hilbert modular form f gives rise to an element of $S(U^\circ)$ which is an eigenform for $\mathbf{T}$, and hence we get a map $\mathbf{T} \to \mathcal{O}$. Set $\mathfrak{m}$ to be the unique maximal ideal of $\mathbf{T}$ containing the kernel of this map.

Now let Q be a Taylor-Wiles set of primes (disjoint from v_{aux}). Let R_Q be what it was in Mike's talk, with the additional caveat that we permit ramification at v_{aux}. Thus

$$R_Q = \widetilde{R} \otimes_{B_0} B$$

where $\widetilde{R}$ is the universal global deformation ring of $\overline{\rho}$ with determinant χ_p, unramified outside $St \cup S_p \cup Q \cup \{v_{aux}\}$; the ring B_0 is the product of the universal local deformation rings with the right determinant, at the places in $St \cup S_p$; the modification B is the product of the universal Steinberg deformation rings at places in St, and "suitably modified" universal ord-cryst deformation rings at places in S_p.

Recall that the universal deformation

$$\rho_Q : G_F \to \text{GL}_2(R_Q)$$

when restricted to G_{F_w} for $w \in Q$, has the form

$$\left(\begin{smallmatrix} \eta_1 & 0 \\ 0 & \eta_2 \end{smallmatrix}\right)$$

for tamely ramified characters η_i. As Mike discussed, via class field theory, this endows R_Q with the structure of an $\mathcal{O}[\Delta_Q]$-algebra; here Δ_Q is the product (over the primes $v \in Q$) of the maximal p-power quotients of the cyclic groups $(\mathcal{O}_{F_v}/\mathfrak{p}_v)^\times$. So

$$\mathcal{O}[\Delta_Q] \approx \frac{\mathcal{O}[[y_1, \ldots, y_h]]}{((y_1+1)^{p^{a_1}} - 1, \ldots, (y_h+1)^{p^{a_h}} - 1)},$$

where $h = \#Q$ and the a_i are integers ≥ 1, and in fact $\geq n$ if $Q = Q_n$ is part of a Taylor-Wiles system. Note that $\mathcal{O}[\Delta_Q]$ is a local ring, with maximal ideal $\langle p, y_1, \ldots, y_h \rangle$, because all the $y_1 + 1$ have p-power order, i.e. because Δ_Q is a p-group. We'll use this fact later.

Moreover this $\mathcal{O}[\Delta_Q]$-algebra structure on R_Q is essentially canonically determined by Q, as long as we include in the data of Q a *choice* of one of the two distinct eigenvalues of $\overline{\rho}(\mathrm{Frob}_v)$ for each $v \in Q$; this lets us pick out one of the characters η_i comprising $G_F \to \mathrm{GL}_2(R_Q)$ to be "η_1", to which we can then apply class field theory and get the map $\Delta_Q \to R_Q^\times$ as Mike explained. Call this distinguished eigenvalue $\alpha_v \in k$.

We let $a_Q \lhd \mathcal{O}[\Delta_Q]$ be the **augmentation ideal** – recall that in any group ring $A[G]$, this is the kernel of the map $A[G] \to A$ which sends each $g \in G$ to 1. This ideal will show up later, in relating the modules $M_n^\square$ to $M_0^\square$. In the presentation above, it is generated by the y_is, one for each element of the TW set Q. For now, let's see how R_Q is related to $R_\varnothing = R'$.

Lemma 2.0.2. The canonical map $R_Q \to R_\varnothing$ is surjective with kernel $a_Q R_Q$.

Proof. We show that $G_F \to \mathrm{GL}_2(R_Q) \to \mathrm{GL}_2(R_Q/a_Q)$ is universal for the appropriate deformation problem. Fix a deformation $\rho_A : G_F \to \mathrm{GL}_2(A)$ of $\overline{\rho}$, which is ordinary-crystalline at places over p, Steinberg at places in St, and ramified only in $St \cup S_p \cup \{v_{\mathrm{aux}}\}$. Then we certainly get a map $\varphi_A : R_Q \to A$ such that $\varphi_A \circ \rho_Q = \rho_A$. But since ρ_A is unramified at any $w \in Q$, when composed with φ_A the distinguished character $\eta_{1,w}$ is trivial on inertia. Thus if Δ_w is the maximal p-power quotient of $(\mathcal{O}_{F_w}/\mathfrak{p}_w)^\times$ and $\delta : \Delta_Q \to R_Q^\times$ the map Mike discussed, we have $\varphi_A \circ \delta(\sigma) = 1$ for all $\sigma \in \Delta_w$. This holds for all $w \in Q$, so the elements $1 - \delta(\sigma)$ for $\sigma \in \Delta_Q$ are all killed by φ_A. These elements generate the augmentation ideal a_Q, so φ_A factors through $R_Q/a_Q R_Q$. But if $\varphi : R_Q/a_Q R_Q \to A$ were another map lifting ρ_A, then the composition of φ with the projection from R_Q would have to agree with φ_A by universality of R_Q. Since said projection is, of course, surjective, this shows that $(R_Q/a_Q R_Q, \rho_Q \bmod a_Q)$ is universal for the type of deformations we want. □

The lemma shows that we know exactly how to relate R_Q to its level zero version $R_\varnothing$. For patching, we now want to set certain R_Q-modules M_Q of automorphic forms, which will be free over $\mathcal{O}[\Delta_Q]$ and related to $M_\varnothing$ in the same manner as the lemma.

For this we need to specify some new compact open subgroups $U_Q \subset V_Q \subset U^\circ$, by shrinking (physically speaking) the level at $w \in Q$. These will all agree except for those $w \in Q$. For $w \in Q$ set $V_{Q,w}$ to be the Iwahori I_w:

$$V_{Q,w} = I_w = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathcal{O}_{F_w}) : c \in \mathfrak{p}_w\}.$$

Set

$$U_{Q,w} = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathcal{O}_{F_w}) : c \in \mathfrak{p}_w, d = a^{-1} \in \Delta_w \right\},$$

i.e. the image of $ad^{-1} \bmod \mathfrak{p}_w$ should map to 1 in the maximal p-power quotient, for all $w \in Q$.

The following is clear.

Lemma 2.0.3. U_Q is normal in V_Q, and $V_Q/U_Q = \Delta_Q$. □

This means that the induced map of sets $X(U_Q) \to X(V_Q)$ "wants to be" a Δ_Q-torsor. The role of the auxiliary prime v_{aux} will be to ensure that this is the case, as will be discussed below. This Galois property of the aforementioned cover will then be used to show that the module M_Q we shall define next, is actually $\mathcal{O}[\Delta_Q]$-free.

Now we have Hecke algebras $\mathbf{T}(V_Q)$ and $\mathbf{T}(U_Q)$, which, as we have defined them, contain only T-operators for places away from $St \cup S_p \cup \{v|\infty\} \cup Q \cup \{v_{aux}\}$ and nothing else. We also want some U-operators for places in Q. So set

$$\mathbf{T}(U_Q)^+ = \langle \mathbf{T}(U_Q), \{U_w : w \in Q\} \rangle \subset \mathrm{End}(S(U_Q)),$$

$$\mathbf{T}(V_Q)^+ = \langle \mathbf{T}(V_Q), \{U_w : w \in Q\} \rangle \subset \mathrm{End}(S(V_Q)).$$

(Note that we've avoided notational ambiguity, since the U_w is distinct from the w-component of the level zero compact open subgroup U_w°. Still, I'm sorry that they look so similar.)

I'll remind you that the U_w operator is given by the Iwahori-double coset

$$I_w \left(\begin{smallmatrix} \varpi_w & 0 \\ 0 & 1 \end{smallmatrix}\right) I_w.$$

Now let us define some ideals in Hecke algebras. As before we set $\mathfrak{m}[= \mathfrak{m}_\varnothing = \mathfrak{m}^\circ]$ to be the maximal ideal of $\mathbf{T} = \mathbf{T}(U^\circ)$ containing the kernel of the eigenvalue map for the automorphic form on $(D \otimes \mathbf{A}_F^f)^\times$ corresponding to our modular form f. There is a natural map $\mathbf{T}(V_Q) \to \mathbf{T}$ sending the T-operators to themselves; set $\mathfrak{m}_Q$ to be the contraction of $\mathfrak{m}$ along this homomorphism. This is a maximal ideal of $\mathbf{T}(V_Q)$. Now set $\mathfrak{n}_Q$ to be the ideal of $\mathbf{T}(V_Q)^+$ generated by $\mathfrak{m}_Q$ plus $U_w - \widetilde{\alpha_w}$ for all $w \in Q$, where $\widetilde{\alpha_w}$ is any lift of $\alpha_w \in k$ to $\mathcal{O} \subset \mathbf{T}(V_Q)^+$. Similarly we have a map $\mathbf{T}(U_Q)^+ \to \mathbf{T}(V_Q)^+$. So we can contract $\mathfrak{n}_Q$ to get an ideal $\mathfrak{m}_Q^+$ of $\mathbf{T}(U_Q)^+$.

Here is the picture:

$$\begin{array}{ccccc} & & \mathfrak{m}_Q \lhd \mathbf{T}(V_Q) & & \mathfrak{m}_Q^+ \lhd \mathbf{T}(U_Q)^+ \\ & \swarrow & & \searrow & \swarrow \\ \mathfrak{m}_\varnothing \lhd \mathbf{T}(V_\varnothing) = \mathbf{T} & & & & \mathfrak{n}_Q \lhd \mathbf{T}(V_Q)^+ \end{array}$$

Now $\mathfrak{n}_Q$ is clearly maximal, provided it is not the unit ideal. (We're just setting all the generators equal to constants.) Why isn't it the unit ideal? In fact, this will come out of what we ultimately prove about $S(V_Q)_{\mathfrak{n}_Q}$, effectively that there is a modular form of level V_Q with action of the Hecke algebra $\mathbf{T}(V_Q)^+$ specified by $\mathfrak{n}_Q$, so the quotient $\mathbf{T}(V_Q)^+/\mathfrak{n}_Q$ is not the zero ring.

Remark 2.0.4. As Andrew explained to me, this can also be seen directly, as follows. (I'm kind of confused about this, and I couldn't really work out the details, but it shouldn't be too important for what follows...) Take a modular form f on D of level U°, hence level 1 at some $w \in Q$. Then f and $\left(\begin{smallmatrix} 1 & 0 \\ 0 & \varpi_w \end{smallmatrix}\right) f$ span the space of forms where we ramp up to Γ_0 (Iwahori) level at w. We need to check that the U_w operator has an eigenvector in this space of forms, with eigenvalue matching our chosen lift $\widetilde{\alpha_w}$. Since f has level 1 at w, the corresponding automorphic representation is unramified principle series at w, i.e. it is of the form $V = \pi(\mu, \nu) = \{\varphi : \mathrm{GL}_2(F_w) \to \mathbf{C} : \varphi((\begin{smallmatrix} a & * \\ & b \end{smallmatrix}) g) = \mu(a)\nu(b)|a/b|_w^{1/2} f(g)\}$ for some unramified characters $\mu, \nu : F_w^\times \to \mathbf{C}^\times$. The dimension of the spherical fixed vectors in V is 1; this space is spanned by f itself Then f and $(\begin{smallmatrix} 1 & \\ & \varpi_w \end{smallmatrix}) f$ span the Iwahori fixed vectors V^{I_w}. The U_w operator is given by the double coset

$$I_w \left(\begin{smallmatrix} \varpi_w & \\ & 1 \end{smallmatrix}\right) I_w = \coprod_{x \in (\mathcal{O}_{F_w}/\mathfrak{p}_w)} \left(\begin{smallmatrix} \varpi_w & \tilde{x} \\ & 1 \end{smallmatrix}\right) I_w.$$

Using this decomposition and the basis of indicator functions for the two cells in

$$G = \mathrm{GL}_2(F_w) = B \sqcup B \left(\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right) I_w, \qquad (B = \text{borel})$$

one can compute the action of U_w explicitly. The eigenvalues should be $\mu(\varpi_w), \nu(\varpi_w)$, one of which should match $\widetilde{\alpha_w}$ somehow (?).

So we have a maximal ideal $\mathfrak{m}_Q^+$ in $\mathbf{T}(U_Q)^+$. We will localize at this to define the rings of Hecke operators T_Q and modules of automorphic forms M_Q which we will use for patching: set

$$T_Q = \mathbf{T}(U_Q)^+_{\mathfrak{m}_Q^+}, \qquad M_Q = S(U_Q)_{\mathfrak{m}_Q^+}.$$

There is a natural Galois representation $G_F \to \mathrm{GL}_2(T_Q)$ which induces a surjection $R_Q \twoheadrightarrow T_Q$, and hence an R_Q-module structure on M_Q (which is naturally a T_Q-module). Since R_Q is an $\mathcal{O}[\Delta_Q]$-algebra, this makes M_Q an $\mathcal{O}[\Delta_Q]$-module as well.

Now we can state our main results.

Theorem 2.0.5. *The module M_Q is $\mathcal{O}[\Delta_Q]$-free. Moreover,*

$$M_Q/\mathfrak{a}_Q M_Q \cong M_\varnothing.$$

This should imply more or less directly that certain quotients of appropriate framed versions $M^\square_{Q_n}$ of the M_{Q_n}s form a "patching datum".

3 Proof of Main Theorem

To prove the first part of the theorem (freeness of M_Q over $\mathcal{O}[\Delta_Q]$), we will argue sort of topologically. We will show that the action on $\mathcal{O}[\Delta_Q]$ on M_Q via the map $\mathcal{O}[\Delta_Q] \to R_Q \to T_Q \subset \mathrm{End}(M_Q)$, agrees with another action which is a bit easier to understand. Specifically, we will show that the map of double coset spaces

$$X(U_Q) \to X(V_Q)$$

is a Galois cover with deck group $\Delta_Q = V_Q/U_Q$, and deduce from this that Δ_Q acts freely on M_Q.

Remark 3.0.6. If you look in DDT for the proof of the analgous freeness assertion in their setup (Thm 4.16), you'll see that they invoke the fact that a certain map of modular curves corresponding to an inclusion of congruence subgroups, is unramified, and they proceed using Riemann-Hurwitz. Our strategy here is thus sort of similar to theirs.

To ensure that said map is a Galois cover, i.e. a Δ_Q-torsor, we need to specify the group level U° at the auxiliary prime ν_{aux} I mentioned earlier. Essentially, by adding enough level at that place, we can ensure that the "stackiness" – the order of the automorphism group $N_{U_Q,x}$ for various $x \in X(U_Q)$ – is prime to p, hence prime to the index $\#\Delta_Q = [V_Q : U_Q]$, for all x. This is what we turn to now.

3.1 The smallness condition and ν_{aux}

First let us relate the $V_Q/U_Q = \Delta_Q$ action on the fibers of $X(U_Q) \to X(V_Q)$, to the amount of stackiness. This requires a bit of group theory.

3.1.1 Some group theory

First note the following trivial fact.

Lemma 3.1.1. Suppose we have subgroups A, B, C of a group G, and suppose $C \lhd B$. Then $(A \cap B)/(A \cap C)$ is naturally a subgroup of B/C via $(A \cap B)/(A \cap C) \hookrightarrow B/(A \cap C) \twoheadrightarrow B/C$. $\square$

Lemma 3.1.2. Suppose we have groups and subgroups

$$K' \lhd K \subset G \supset H.$$

Consider the obvious map of double coset spaces

$$\pi : H\backslash G/K' \twoheadrightarrow H\backslash G/K.$$

Let $S(g_0)$ be the fiber $\pi^{-1}(Hg_0K)$. Set $J(g_0) = (K \cap g_0^{-1}Hg_0)/(K' \cap g_0^{-1}Hg_0)$. By the previous lemma we can regard $J(g_0)$ as a subgroup of K/K'. Then $S(g_0)$ is naturally in bijection with the left coset space $J(g_0)\backslash(K/K')$.

Proof. Define $\alpha : J(g_0)\backslash(K/K') \to S(g_0)$ by

$$\alpha : J(g_0) \cdot (kK') \mapsto Hg_0kK'.$$

We can see α is a well-defined map of sets as follows. If $J(g_0) \cdot (k_1K') = J(g_0) \cdot (k_2K')$ then there exists $j \in J \subset K/K'$ such that $j(k_1K') = k_2K'$. Now $j = k(K' \cap g_0^{-1}Hg_0)$ for some $k \in K \cap g_0^{-1}Hg_0$. Say $k = g_0^{-1}hg_0$. The map of the previous lemma regards j as an element of K/K' as the coset kK'. The condition $j(k_1K') = k_2K'$ says that $kk_1K' = k_2K'$. So $k_2 = kk_1k'$ for some $k' \in K'$. Now $\alpha(J(g_0) \cdot (k_1K')) = Hg_0k_1K'$ while $\alpha(J(g_0) \cdot (k_2K')) = Hg_0k_2K'$. So we must show these agree. But $Hg_0k_2K' = Hg_0kk_1k'K' = Hg_0(g_0^{-1}hg_0)k_1k')K' = Hhg_0k_1k'K' = Hg_0k_1K'$, so they do.

Conversely, define $\beta : S(g_0) \to J(g_0)\backslash(K/K')$ by $Hg_0kK' \mapsto J(g_0)(kK')$. Again, we must check this is well-defined. If $Hg_0k_1K' = Hg_0k_2K'$ then $g_0k_1 = hg_0k_2k'$ for some

$h \in H, k' \in K'$. so $k_1 = g_0^{-1} h g_0 k_2 k'$. In particular this means $g_0^{-1} h g_0 \in K \cap g_0^{-1} H g_0$. Call this element k. So we need to show that $J(g_0)(k_2 K') = J(g_0)(k k_2 k' K')$. Clearly we can ignore the k'. We leave to the reader to check that $k k_2 K' = (k(K' \cap g_0^{-1} H g_0)) \cdot (k_2 K')$ when we regard $k(K' \cap g_0^{-1} H g_0) \in J$ as an element of K/K'. So $k k_2 k' K' = k k_2 K'$ is a translate of $k_2 K'$ on the left by an element of $J(g_0)$, so we win.

Finally, it is clear that α and β are mutually inverse bijections. □

Remark 3.1.3. Note that the lemma identifies $S(g_0)$ with $J(g_0)\backslash(K/K')$ note merely as sets, but as (K/K')-sets. Hence to prove that $H\backslash G/K' \to H\backslash G/K$ is a K/K'-torsor, it suffices to ensure that $J(g_0)$ vanishes for all $g_0 \in G$.

3.1.2 Application to our setup

In our setup, the previous lemma identifies the fiber of $X(U_Q) \to X(V_Q)$ over $D^\times x V_Q(\mathbf{A}_F^f)^\times$ as a Δ_Q-set with the quotient

$$\frac{(V_Q \cdot (\mathbf{A}_F^f)^\times)/(U_Q \cdot (\mathbf{A}_F^f)^\times)}{(x^{-1} D^\times x \cap V_Q \cdot (\mathbf{A}_F^f)^\times)/(x^{-1} D^\times x \cap U_Q \cdot (\mathbf{A}_F^f)^\times)}.$$

(We think of this quotient as a space of left-cosets of the denominator.) We'd like to show the denominator is trivial, so that this is simply the quotient

$$V_Q(\mathbf{A}_F^f)^\times / U_Q(\mathbf{A}_F^f)^\times = V_Q/U_Q = \Delta_Q,$$

so the action on fibers is simply transitive as required.

Now the denominator is itself a quotient of

$$\frac{x^{-1} D^\times x \cap V_Q(\mathbf{A}_F^f)^\times}{x^{-1} D^\times x \cap (\mathbf{A}_F^f)^\times}.$$

The denominator of the latter is simply $F^\times$.

So we want to ensure that

$$(x^{-1} D^\times x \cap V_Q(\mathbf{A}_F^f)^\times)/F^\times = \{1\}.$$

Since $V_Q \subset U^\circ$, we can simply impose conditions on the ground level U° so that

$$\mathcal{G} := (x^{-1} D^\times x \cap U^\circ(\mathbf{A}_F^f)^\times)/F^\times = \{1\}.$$

Now $x^{-1} D^\times x$ is discrete, and $U^\circ(\mathbf{A}_F^f)^\times/F^\times$ is compact. (Because U° is compact, and the finite part of the idéle class group is compact.) So this is a finite group.

Now choose ν_{aux} lying over some prime $\ell_{aux} \geq 5$, sufficiently large so that ν_{aux} is unramified for both F and D (and is outside $St \cup S_p$). Set

$$U^\circ_{\nu_{aux}} = \{m \in GL_2(\mathcal{O}_{F_{\nu_{aux}}}) : m \equiv 1 \bmod \mathfrak{p}_{\nu_{aux}}\}.$$

Observe that $U^\circ_{\nu_{aux}}$ is pro-ℓ_{aux}.

Proposition 3.1.4. $GL_2(F_{\nu_{aux}})$ has no elements of order ℓ_{aux}.

This is because $F_{v_{aux}}$ is an unramified extension of $\mathbf{Q}_{\ell_{aux}}$, so it does not contain ℓ_{aux} roots of unity. But if $M \in GL_2(F_{v_{aux}})$ had order ℓ_{aux}, it would have an ℓ_{aux} root of unity as an eigenvalue, lying in some quadratic extension of $F_{v_{aux}}$. But for $\ell_{aux} \geq 5$, $[F_{v_{aux}}(\mu_{\ell_x}) : F_{v_{aux}}]$ cannot be two.

By the proposition, it follows that $D^\times$ has no elements of ℓ_{aux}-power order, since $D^\times \hookrightarrow (D \otimes_F F_{v_{aux}})^\times \approx GL_2(F_{v_{aux}})$ as v_{aux} splits D. Since the finite group $\mathcal{G}$ we want to show is trivial is a subquotient of (a conjugate of) $D^\times$, we know that its order is thus prime to ℓ_{aux}.

To prove it is trivial, we will show our group $\mathcal{G}$ is an ℓ_{aux}-group.

So suppose $\overline{g} \in \mathcal{G}$ had order n prime to ℓ_{aux}. Fix a representative $g \in x^{-1}D^\times x \cap U^\circ(\mathbf{A}_F^f)^\times$. Consider the v_{aux} component $g_{v_{aux}}$. Inside $(D \otimes_F F_{v_{aux}})^\times$, $g_{v_{aux}}$ sits in the subgroup

$$x_{v_{aux}}^{-1} D^\times x_{v_{aux}} \cap U^\circ_{v_{aux}} F^\times_{v_{aux}} \subset (D \otimes F_{v_{aux}})^\times \approx GL_2(F_{v_{aux}}).$$

Write $g_{v_{aux}} = uj$ as product of something in $U^\circ_{v_{aux}}$ and someting in $F^\times_{v_{aux}}$ (here j stands for "idele") Now $(uj)^n = u^n j^n \in F^\times$, so $u^n \in F^\times_{v_{aux}}$. Thus the image of u in $U^\circ_{v_{aux}}/U^\circ_{v_{aux}} \cap F^\times_{v_{aux}}$ has order prime to ℓ_{aux}. But this group is pro-ℓ_{aux}, being a quotient of $U^\circ_{v_{aux}}$, and hence $u = 1$. Thus $g_{v_{aux}} \in F^\times_{v_{aux}} \cap x^{-1}_{v_{aux}} D^\times x_{v_{aux}}$. In particular, $g_{v_{aux}}$ commutes with $x_{v_{aux}}$, so $g_{v_{aux}} \in F^\times_{v_{aux}} \cap D^\times = F^\times$. This shows that in fact $g \in F^\times$ so $\overline{g} = 1 \in \mathcal{G}$ as desired.

3.2 Proof of freeness

OK great, so now we have seen that $X(U_Q) \to X(V_Q)$ is a Δ_Q-torsor, provided we impose "principal congruence subgroup" level in U° at a well-chosen place v_{aux}. In particular, this shows that $X(U_Q)$ is (non-canonically) the same as $X(V_Q) \times \Delta_Q$. So $S(U_Q) = S(V_Q) \otimes_{\mathcal{O}} \mathcal{O}[\Delta_Q]$. Since $S(V_Q)$ is $\mathcal{O}$-finite free, this means that $S(U_Q)$ is $\mathcal{O}[\Delta_Q]$-finite free. (That is, with the action of $\mathcal{O}[\Delta_Q]$ as deck transformations of $X(U_Q)$.)

Now the localization $M_Q = S(U_Q)_{\mathfrak{m}_Q^+}$ is a summand of $S(U_Q)$ as a T_Q-module. (This is because T_Q is finite semilocal over the $\mathfrak{p}$-adically complete ring $\mathcal{O}$.) So provided we know that the deck transformation action of $\mathcal{O}[\Delta_Q]$ agrees with the action coming from the homomorphism

$$\mathcal{O}[\Delta_Q] \to R_Q \to T_Q \subset \mathrm{End}(M_Q)$$

this shows that that M_Q is a summand of a finite free $\mathcal{O}[\Delta_Q]$-module. Since $\mathcal{O}[\Delta_Q]$ is local, that would force M_Q to be $\mathcal{O}[\Delta_Q]$-free. (Projective = flat = locally free = free, for finitely generated modules over a local Noetherian ring.)

Thus we are reduced to showing that these two actions of the "diamond operators" Δ_Q agree. For this we will need to know the following.

Lemma 3.2.1. The Hecke algebra T_Q is reduced.

Proof. (FIXME: Cf. Taylor Cor. 1.8(3) in "On the Meromorphic Continuation...")

The rough idea is the following. It suffices to consider the generic fiber of the Hecke algebra, i.e. to consider the space of forms after inverting $\mathfrak{p}$ in our coefficients. Reducedness of the Hecke algebra says that eigenforms are determined by their Hecke eigenvalues at all but finitely many places. This is because we defined the Hecke ring as subring of the endomorphisms of the module of module forms; consequently the modular forms are faithful over

the Hecke algebra. Thus if some local quotient of the Hecke algebra is not a field, the corresponding quotient module of modular forms will have positive dimension. This translates directly to the existence of several eigenforms with exactly the same Hecke eigenvalues for almost all places.

So to prove reducedness, it's enough to check that in our situation, the Hecke eigenvalues are actually determined by the corresponding Galois representation which we know (huh?!?), which is determined by the knowledge of what's happening at all but finitely many places. □

By the lemma, the module M_Q is spanned by Hecke eigenforms $f \in S(U_Q)_{\mathfrak{m}_Q^+} = M_Q$, corresponding to the various irreducible components $T_Q/\mathfrak{p}$ ($\mathfrak{p}$ a minimal prime) of T_Q. So for each f (i.e. for each minimal $\mathfrak{p}$) we have two actions of Δ_Q on $M_Q/\mathfrak{p}M_Q$:

- The one coming from the $\Delta_Q = V_Q/U_Q$ action on $X(U_Q)$;
- and the one coming from the morphisms

$$\mathcal{O}[\Delta_Q] \to R_Q \to T_Q \to T_Q/\mathfrak{p}.$$

If we know these agree, for each f, then it follows that the two actions of Δ_Q on all of M_Q agree. So fix such an f. Let π be the representation of $(D \otimes_F \mathbf{A}_F^f)^\times$ generated by f, after tensoring with $E = \mathrm{Frac}(\mathcal{O})$. Fix one of the TW primes $w \in Q$, and consider the local component π_w of π at w. Now f itself is a $U_{Q,w}$-fixed vector. (I'll remind you that $U_{Q,w} = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathcal{O}_{F_w}) : c \equiv 0 \bmod \varpi_w, ad^{-1} = 1 \in \Delta_Q\}$. This is like $\Gamma_1(w)$, sort of.) Moreover we know how the Hecke operator U_w (sorry for the notation!) acts on f: as a lift of one of the eigenvalues of $\overline{\rho}(\mathrm{Frob}_w)$, which are distinct, and whose product is equal to $Nw \equiv 1 \bmod p$, because w is a Taylor-Wiles prime. These stringent conditions on the action of U_w rule out the possibility that π_w is Steinberg, but I am not sure why. Now by a classification result, this implies that π_w is in fact a tamely ramified principal series

$$\pi_w = \pi(\mu, \nu), \qquad \mu, \nu : F_v^\times \to \mathbf{C}^\times \text{ tame.}$$

Now $\Delta_Q = V_Q/U_Q$ acts on the right (in the way we like) on

$$D^\times \backslash (D \otimes \mathbf{A}_F^f)^\times$$

by translation by a representative for the coset $\nu U_Q \in \Delta_Q$. This induces an action of the w-part Δ_w = maximal p-power quotient of $(\mathcal{O}_{F_w}/\varpi_w)^\times$ of Δ_Q, on the invariants $\pi_w^{U_{Q,w}}$. Moreover, if we write

$$\pi_w = \pi(\mu, \nu) = \{\varphi : \mathrm{GL}_2(F_w) \to \mathbf{C} \mid \varphi(\left(\begin{smallmatrix} a & * \\ & b \end{smallmatrix}\right) g) = \mu(a)\nu(b)|a/b|^{1/2}\varphi(g)\}$$

as a space of functions on $D_w^\times$, this action of Δ_w agrees by definition with the usual (= right regular) action of

$$\{\langle x \rangle = \left(\begin{smallmatrix} \widetilde{x} & \\ & 1 \end{smallmatrix}\right) : x \in \Delta_w, \widetilde{x} \text{ a lift to } \mathcal{O}_{F_w}\} \subset \{\langle x \rangle : x \in (\mathcal{O}_{F_w}/\varpi_w)^\times\}.$$

(We really only care about these diamond operators for $x \in \Delta_w$, but they make perfect sense for any $x \in (\mathcal{O}_{F_w}/\varpi_w)^\times$.)

"The following lemma is well-known":

Lemma 3.2.2 (Similar to Taylor, "On the meromorphic continuation...", Lemma 1.6)**.** If μ and ν are tame, then $\pi(\mu,\nu)^{U_{Q,w}}$ is two dimensional, with a basis e_μ, e_ν of U_w-eigenvectors, such that

$$U_w e_\mu = \mu(\varpi_w)e_\mu, \qquad U_w e_\nu = \nu(\varpi_w)e_\nu,$$ [1]

and the diamond operators act by

$$\langle x\rangle e_\mu = \mu(\widetilde{x})e_\mu, \qquad \langle x\rangle e_\nu = \nu(\widetilde{x})e_\nu, \qquad (x \in (\mathcal{O}_{F_w}/\varpi_w)^\times).$$

(Does anyone know a real reference? Probably it's not too hard prove; it should just be some explicit computation.) □

Now f is a U_w-eigenvector with eigenvalue $\widetilde{\alpha_w}$, by the way we set things up, so we can say that μ is determined by α_w plus the action of Δ_w on $\pi_w^{U_{Q,w}}$.

Local Langlands implies that

$$\rho_{f,p} : G_F \to GL_2(\overline{\mathbf{Q}}_p)$$

corresponding to $\pi = \pi_f$, satisfies

$$\rho_{f,p}|_{G_{F_w}} \sim \begin{pmatrix} \mu & \\ & \nu \end{pmatrix}$$

where μ, ν are the Galois characters corresponding to μ, ν via class field theory. [2] But this representation is precisely the one we know as

$$G_{F_w} \to GL_2(R_Q) \to GL_2(T_Q) \to GL_2(T_Q/\mathfrak{p}) \to GL_2(\overline{\mathbf{Q}}_p).$$

So the action of the diamond operators $\langle x\rangle$ for $x \in \Delta_w$ act on f by the values of the character $\mu(\widetilde{x})$, agrees with action of x via the value of the character "$\eta_1(\widetilde{x})$" we picked out when originally defining $\mathcal{O}[\Delta_Q] \to R_Q$.

This essentially proves the desired compatibility between the two actions, modulo all the details I've omitted or gotten wrong. Therefore we have completed the proof of the freeness of M_Q over $\mathcal{O}[\Delta_Q]$.

3.3 Proof of relation of level Q with level ∅

The remaining part of the theorem is the "moreover", namely:

$$M_Q/\mathfrak{a}_Q M_Q \cong M_\varnothing.$$

The key ingredient in the proof of the "moreover" will be the following.

Proposition 3.3.1. There is an isomorphism

$$M_\varnothing := S(U^\circ)_\mathfrak{m} \xrightarrow{\sim} S(V_Q)_{\mathfrak{n}_Q}.$$

[1] This may not be quite right...

[2] Clearly I've been sloppy somewhere regarding $\overline{\mathbf{Q}}_p$ and $\mathbf{C}$....

Assuming the proposition, let us deduce what we want. By the proposition, it is sufficient to prove that

$$M_Q/\mathfrak{a}_Q M_Q \cong S(V_Q)_{\mathfrak{n}_Q},$$

since the latter is the same as $M_\varnothing$. Since we've already "shown" that the two Δ_Q actions are the same, we may as well compute the Δ_Q-coinvariants of M_Q, namely $M_Q/\mathfrak{a}_Q M_Q$, using the description of M_Q as $\mathcal{O}$-valued functions "upstairs" in the Δ_Q torsor

$$X(U_Q) \overset{\Delta_Q}{\twoheadrightarrow} X(V_Q).$$

But with this description, it is more or less obvious that the desired isomorphism holds: we have (non-canonically) that

$$S(U_Q) = S(V_Q) \otimes_{\mathcal{O}} \mathcal{O}[\Delta_Q]$$

so

$$S(U_Q)_{\Delta_Q} = S(V_Q).$$

Since quotients commute with localization, the same equality holds when we localize at $\mathfrak{m}_Q^+$, resp. $\mathfrak{n}_Q$.

It remains only to prove the relationship between $S(V_Q)_{\mathfrak{n}_Q}$ and $S(U^\circ)_{\mathfrak{m}}$.

4 Proof of Proposition

We will prove the proposition by induction on the size of the TW-set Q, reducing to the case when $Q = \{w\}$ is a singleton.

This argument is due to Andrew, the basic outline being from Taylor's "On the meromorphic continuation..." paper (Lemma 2.2).

4.1 Inductive setup

Specifically, let V be any compact open subgroup of U°. Let $\Phi : \mathbf{T}(V) \to k$ be a homomorphism with kernel $\mathfrak{m}$, and let $\widetilde{\Phi} : \mathbf{T}(V) \to \mathcal{O}$ be any set-theoretic lift. Let w be a TW prime, meaning the following.

a) $w \notin \Sigma(V)$ (which, recall, is just the bad set of places: $S_p \cup St \cup \{v|\infty\} \cup \{v : V_v \text{ nonmaximal}\}$). In practice, i.e. for our inductive argument, this means w is nonarchimedean and outside $S_p \cup St$ and the TW primes we already added.

b) $\mathbf{N}w = 1 \bmod p$.

c) $X^2 - \Phi(T_w)X + \mathbf{N}w$ has distinct roots $\alpha, \beta \in k$.

By Hensel's lemma, we obtain a factorization in $\mathbf{T}(V)_{\mathfrak{m}}[X]$:

$$X^2 - T_w X + \mathbf{N}w = (X - A)(X - B)$$

where $\Phi(A) = \alpha, \Phi(B) = \beta$.

Now let V' be the compact open subgroup of V obtained by replacing $V_w = \mathrm{GL}_2(\mathcal{O}_{F_w})$ with $V'_w = I_w$ the Iwahori. Let U_w be the Hecke operator on $S(V')$ given by $I_w \left(\begin{smallmatrix} \varpi_w & \\ & 1 \end{smallmatrix}\right) I_w$ as before. Let $\mathbf{T}(V')^+$ be the subalgebra of $\mathrm{End}(S(V'))$ generated by $\mathbf{T}(V')$ and U_w. Let $\mathfrak{m}'$ be the ideal of $\mathbf{T}(V')^+$ generated by $\mathfrak{p}, T_v - \widetilde{\Phi}(T_v)$ for $v \notin \Sigma(V')$, and $U_w - \widetilde{\alpha}$.

Proposition 4.1.1 (Induction step)**.** There is an isomorphism

$$\eta : S(V)_{\mathfrak{m}} \to S(V')_{\mathfrak{m}'}$$

given by

$$f \mapsto Af - \left(\begin{smallmatrix} 1 & \\ & \varpi_w \end{smallmatrix}\right) f.$$

Granting this, the proof of the main theorem is complete. For using this induction step, we can build a chain of isomorphisms from $S(U^\circ)_{\mathfrak{m}}$ to $S(V_Q)_{\mathfrak{m}_Q^+}$ by adding the Taylow-Wiles primes $w \in Q$ one at a time, invoking the induction step each time.

4.2 Proof of induction step

4.2.1 Well-definedness of η

Note that *a priori* it is not clear that η lands in the localization $S(V')_{\mathfrak{m}'}$. (Recall that $\mathbf{T}(V')_{\mathfrak{m}'}$ is semilocal and finite over $\mathcal{O}$, hence a direct sum of its localizations at its maximal ideals, so in particular we can regard those localizations as subs rather than quotients.) We can characterize $S(V')_{\mathfrak{m}'}$ as precisely the $\mathbf{T}(V')^+$-submodule of $S(V')$ on which $\mathfrak{m}'$ acts topologically nilpotently. (Also, $S(V)_{\mathfrak{m}} \subset S(V)$ is characterized similarly.) As a first step, let us use this characterization to show that η actually lands where we want it to.

The following is a consequence of explicit computations done with double cosets.

Lemma 4.2.1. The identities

$$T_w f = U_w f + \left(\begin{smallmatrix} 1 & \\ & \varpi_w \end{smallmatrix}\right) f, \qquad U_w \left(\begin{smallmatrix} 1 & \\ & \varpi_w \end{smallmatrix}\right) f = \mathbf{N}w \cdot f$$

hold for any $f \in S(V)$. □

As a consequence we have

Lemma 4.2.2. $U_w \circ \eta = \eta \circ A$.

Proof. Using the previous lemma, we can expand

$$U_w \eta(f) = T_w(Af) - \left(\begin{smallmatrix} 1 & \\ & \varpi_w \end{smallmatrix}\right)(Af) - \mathbf{N}w \cdot f.$$

Since A is a root of $X^2 - T_w X + \mathbf{N}w \in \mathbf{T}(V)_{\mathfrak{m}}[X]$, we have

$$A^2 = T_w A - \mathbf{N}w.$$

Hence

$$\eta(Af) = A^2 f - \left(\begin{smallmatrix} 1 & \\ & \varpi_w \end{smallmatrix}\right)(Af) = T_w A - \mathbf{N}w \cdot f - \left(\begin{smallmatrix} 1 & \\ & \varpi_w \end{smallmatrix}\right)(Af) = U_w \eta(f).$$ □

Now we deduce that η is well-defined. We need to show that p, $T_\nu - \widetilde{\Phi}(T_\nu)$ for $\nu \notin \Sigma(V')$, and $U_w - \widetilde{\alpha}$ all act topologically nilpotently on $\eta(f)$ for any $f \in S(V)_m$.

For p this is clear.

For $T_\nu - \widetilde{\Phi}(T_\nu)$, one checks that T_ν commutes with η. It is clear that T_ν commutes with A, since $\mathbf{T}(V)$ commutes with the right regular action of $\left(\begin{smallmatrix} 1 & \\ & \varpi_w \end{smallmatrix}\right)$; that follows from a calculation with the appropriate double coset. Consequently the topological nilpotence of $T_\nu - \widetilde{\Phi}(T_\nu)$ on $\eta(f)$ follows from the topological nilpotence of the same operator acting on $f \in S(V)_m$.

Finally, by the previous lemma we have $(U_w - \widetilde{\alpha})(\eta f) = \eta(A - \widetilde{\alpha})(f)$. But this is topologically nilpotent since f is in $S(V)_m$, hence $A - \widetilde{\alpha}$, acts topologically nilpotently on f.

[FIXME: Explain the last sentence.]

4.2.2 Aside: the "integration pairing" on $X(U)$

Next we will show that η is injective. To do so, we will make use of an "integration pairing"

$$\langle , \rangle_U : S(U) \otimes S(U) \to \mathcal{O}$$

for any $U \subset U^\circ$. This is defined by

$$\langle f, g \rangle_U = \sum_{x \in X(U)} f(x) g(x).$$

It is just the "L^2 inner product" with respect to the counting measure on $X(U)$.

As I think Akshay mentioned several lectures ago, in principle we should use a different measure: we should weight a point x by the amount of "stackiness" of $X(U)$ at x:

$$N_{U,x} = [x^{-1} D^\times x \cap U \cdot (\mathbf{A}_F^f)^\times : F^\times].$$

But we arranged U° so that $N_{U,x}$ is automatically 1.

4.2.3 Injectivity

Lemma 4.2.3. For $f, g \in S(V)$, we have

$$\langle f, T_w g \rangle_V = \langle f, \left(\begin{smallmatrix} 1 & \\ & \varpi_w \end{smallmatrix}\right) g \rangle_{V'}.$$

Proof. An explicit computation with double cosets, which we omit. □

Lemma 4.2.4. Let $\pi : S(V') \to S(V)$ be the adjoint to the inclusion $S(V) \hookrightarrow S(V')$. Then the composition

$$S(V)_m \xrightarrow{\eta} S(V')_{m'} \xrightarrow{\pi} S(V)_m$$

equals $\mathbf{N}w \cdot A - B$.

Proof. A similar argument to what we did above using topological nilpotence shows that the composition above is well-defined, i.e. lands in $S(V)_m$. It uses the fact that the adjoint π respects the action of the T_νs.

Now fix $g \in S(V)$. The adjoint $\pi(\eta(g))$ is characterized by

$$\langle f, \pi(\eta(g))\rangle_V = \langle f, \eta(g)\rangle_{V'}, \qquad \forall f \in S(V).$$

So

$$\langle f, \pi\eta(g)\rangle_V = \langle f, \eta g\rangle_{V'} = \langle f, Ag - \left(\begin{smallmatrix} 1 & \\ & \varpi_w \end{smallmatrix}\right) g\rangle_{V'}.$$

Now

$$\langle f, Ag\rangle_{V'} = \sum_{x \in X(V')} f(x)(Ag)(x) = \sum_{y \in X(V)} \sum_{X(V') \ni x \mapsto y} f(x)(Ag)(x)$$

$$= [V : V'] \sum_{y \in X(V)} f(y)(Ag)(y) = [V : V']\langle f, Ag\rangle_V.$$

On the other hand by the last lemma

$$\langle f, \left(\begin{smallmatrix} 1 & \\ & \varpi_w \end{smallmatrix}\right) g\rangle_{V'} = \langle f, T_w g\rangle_V.$$

So we see

$$\langle f, \pi\eta g\rangle_V = \langle f, [U : U']Ag - T_w g\rangle_U.$$

This shows that $\pi\eta g = ([U : U']A - T_w)g$. But $[U : U'] = \#\mathbf{P}^1(\Bbbk(w)) = \mathbf{N}w + 1$, so $\pi\eta g = (\mathbf{N}w \cdot A + A - (A + B))g = (\mathbf{N}wA - B)g$. □

To conclude that η is injective, by the last lemma it suffices to show that $\mathbf{N}w \cdot A - B$ is a unit. We just need to show it is not in $\mathfrak{m}$. But $\mathbf{N}w = 1 \bmod \mathfrak{p}$, so

$$\mathbf{N}w \cdot A - B \equiv \alpha - \beta \bmod \mathfrak{m},$$

and this is nonzero because α and β were assumed distinct.

So we crucially used the fact that we are at a TW-prime!

4.2.4 Cokernel is torsion-free

In fact, the last proof gives us a bit more. If we scale π by the inver of $\mathbf{N}wA - B$, we get a genuine section of η. So the image of η is a summand of $S(V')_{\mathfrak{m}'}$, and hence the cokernel of η is torsion free.

4.2.5 Surjectivity

It remains to prove that η is surjective.

The first key point is to show that $S(V')_{\mathfrak{m}'}$ (and hence the image of η) is contained in the space of old-forms, i.e. those coming from $S(V)$ either by the inclusion or by $f \mapsto \left(\begin{smallmatrix} 1 & \\ & \varpi_w \end{smallmatrix}\right) f$. Call this space $\mathrm{Old}(V')$.

Lemma 4.2.5. $S(V')_{\mathfrak{m}'} \subset \mathrm{Old}(V')$.

Proof. This is a representation theoretic argument, which I sketch, following Andrew's writeup of it.

Fix $f \in S(V')_{\mathfrak{m}'}$. After tensoring with $E = \mathrm{Frac}(\mathcal{O})$, f generates an automorphic representation of $(D \otimes \mathbf{A}_F^f)^\times$, perhaps reducible, say $\pi = \bigoplus \pi^{(i)}$. We can decompose $f = \sum f^{(i)}$ according to this decomposition. The component $f^{(i)}$ is an I_w-fixed vector in $\pi_w^{(i)}$, for each i. By the representation theory of $GL_2(F_w)$, this forces $\pi_w^{(i)}$ to be an irreducible unramified principle series, or a "special representation with trivial central character".

The latter case will be ruled out the fact that since $f \in S(V')_{\mathfrak{m}'}$, we know how the U_w operator acts on it. Indeed, U_w acts on the Iwahori invariants $(\pi_w^{(i)})^{I_w}$ of a special representation $\pi_w^{(i)}$ by ± 1. In the special case, the Iwahori fixed vectors are 1-dimensional, so they are spanned by $f^{(i)}$. So U_w acts on $f^{(i)}$ by ± 1. Thus $U_w \pm 1$ acts topologicaly nilpotently on $f^{(i)}$. Now $\alpha\beta = Nw = 1 \bmod \mathfrak{p}$. Since $\alpha \neq \beta$, it cannot be the case that $\alpha = \pm 1 \in k$. So since $U_w \pm 1$ acts topologically nilpotently on $f^{(i)}$, $U_w - \widetilde{\alpha}$ cannot, as the difference $\widetilde{\alpha} \pm 1$ acts invertibly. This holds for each i, so $U_w - \widetilde{\alpha}$ cannot act topologically nilpotently on f. But this contradicts the fact that $f \in S(U')_{\mathfrak{m}'}$.

It follows that each $\pi_w^{(i)}$ is an irreducible unramified principle series. This means the spherical fixed vectors are one-dimensional – spanned by sph, say – and the Iwahori fixed vectors are spanned by

$$\mathrm{sph}, \qquad \begin{pmatrix} 1 & \\ & \varpi_w \end{pmatrix} \cdot \mathrm{sph}.$$

These are all oldforms. In particular, the Iwahori fixed vector $f^{(i)}$ is in $\mathrm{Old}(V')$. So the linear combination $f = \sum f^{(i)}$ is in $\mathrm{Old}(V')$ too. □

But in fact we want more: we want $\mathrm{im}(\eta)$ to be in the old-forms *coming from* $S(V)_{\mathfrak{m}}$!. Call this space $\mathrm{Old}_{\mathfrak{m}}(V')$. It's now convenient to introduce some notation: write F for the level-raising map

$$S(V) \oplus S(V) \to S(V')$$
$$(f_1, f_2) \mapsto f_1 + \begin{pmatrix} 1 & \\ & \varpi_w \end{pmatrix} f_2.$$

Lemma 4.2.6. $S(V')_{\mathfrak{m}'} \subset \mathrm{Old}_{\mathfrak{m}}(V') = F(S(V)_{\mathfrak{m}}^{\oplus 2})$.

Proof. F respects the action of $\mathbf{T}(V')$. This implies that for any maximal ideal $\mathfrak{n}$ of $\mathbf{T}(V')$, we have $F(S(V)_{\mathfrak{n}}^{\oplus 2}) \subset S(V')_{\mathfrak{n}}$ and $F^{-1}(S(V')_{\mathfrak{n}}) = S(V)_{\mathfrak{n}}^{\oplus 2}$.

Apply this with $\mathfrak{n} = \mathfrak{m}_0$, the maximal ideal of $\mathbf{T}(V')$ generated by the $T_v - \widetilde{\Phi}(T_v)$ for $v \notin \Sigma(V')$. By definition, $S(V')_{\mathfrak{m}'} \subset S(V')_{\mathfrak{m}_0}$. By the last lemma,

$$F(S(V)^{\oplus 2}) \supset S(V')_{\mathfrak{m}'},$$

so

$$S(V')_{\mathfrak{m}'} \subset S(V')_{\mathfrak{m}_0} \subset F(F^{-1}(S(V')_{\mathfrak{m}_0})) = F(S(V)_{\mathfrak{m}_0}^{\oplus 2}).$$

But by multiplicity one, $S(V)_{\mathfrak{m}_0} = S(V)_{\mathfrak{m}}$, since the only difference between $\mathfrak{m}_0$ and $\mathfrak{m}$ is the presence of the single Hecke operator $T_w - \widetilde{\Phi}(T_w)$. □

An easy computation shows:

Lemma 4.2.7. $U_w F \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = F \begin{pmatrix} T & Nw \\ -1 & 0 \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$. □

Lemma 4.2.8. $S(V)_{\mathfrak{m}}^{\oplus 2}$ is the direct sum of submodules $\mathcal{A} = \{(Af, -f) : f \in S(U)_{\mathfrak{m}}\}$ and $\mathcal{B} = \{(Bf, -f) : f \in S(U)_{\mathfrak{m}}\}$. Moreover we have

$$\left(\begin{smallmatrix} T_w & N_w \\ -1 & 0 \end{smallmatrix}\right) a = Aa, \qquad a \in \mathcal{A}$$

$$\left(\begin{smallmatrix} T_w & N_w \\ -1 & \end{smallmatrix}\right) b = Bb, \qquad b \in \mathcal{B}.$$

Proof. We compute

$$T_w Af - N w f = A^2 f$$

by a previous calculation. So

$$\left(\begin{smallmatrix} T_w & Nw \\ -1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} Af \\ -f \end{smallmatrix}\right) = \left(\begin{smallmatrix} A^2 f \\ -Af \end{smallmatrix}\right) = A\left(\begin{smallmatrix} Af \\ -f \end{smallmatrix}\right).$$

The computation for B is similar. The decomposition $S(V)_{\mathfrak{m}}^{\oplus 2} = \mathcal{A} \oplus \mathcal{B}$ is true because the determinant of the "change of basis matrix" is

$$\det\left(\begin{smallmatrix} A & B \\ -1 & -1 \end{smallmatrix}\right) = A - B \equiv \alpha - \beta \neq 0 \bmod \mathfrak{m}$$

which is a unit in $\mathbf{T}(V)_{\mathfrak{m}}$. □

Lemma 4.2.9. F restricts to a surjection $\mathcal{A} \twoheadrightarrow S(V')_{\mathfrak{m}'}$.

Proof. For $a \in \mathcal{A}, b \in \mathcal{B}$, the last lemmas show

$$U_w F(a+b) = F\left(\begin{smallmatrix} T_w & Nw \\ -1 & 0 \end{smallmatrix}\right)(a+b) = F(Aa + Bb).$$

So

$$(U_w - \widetilde{\alpha})F(a+b) = F((A-\widetilde{\alpha})a + (B-\widetilde{\alpha})b).$$

Iterating this gives

$$(U_w - \widetilde{\alpha})^n F(a+b) = F((A-\widetilde{\alpha})^n a + (B-\widetilde{\alpha})^n b).$$

As $n \to \infty$ this goes to $(B-\widetilde{\alpha})^n F(x')$, since $A - \widetilde{\alpha}$ is topologically nilpotent on $S(V)_{\mathfrak{m}}$. But $B - \widetilde{\alpha}$ is invertible. Consequently if $U_w - \widetilde{\alpha}$ is topologically nilpotent on $F(a+b)$, then $F(b) = 0$. We know that $S(V')_{\mathfrak{m}'} \subset F(\mathcal{A} \oplus \mathcal{B})$, so this implies that in fact $S(V')_{\mathfrak{m}'} \subset F(\mathcal{A})$. □

Finally we can prove the surjectivity of η. For $\eta(f) = F(a)$ where $a = (Af, -f) \in S(V)_{\mathfrak{m}}^{\oplus 2}$. Since everything in $S(V')_{\mathfrak{m}'}$ is of the form $F(a)$ for some a, by the last lemma, and since every $a \in \mathcal{A}$ is of the form $(Af, -f)$ for some $f \in S(V)_{\mathfrak{m}}$, we are done.

POTENTIAL MODULARITY AND APPLICATIONS

ANDREW SNOWDEN

Contents

1. Introduction

In our seminar we have been working towards a modularity lifting theorem. Recall that such a theorem allows one (under suitable hypotheses) to deduce the modularity of a p-adic Galois representation from that of the corresponding mod p representation. This is a wonderful theorem, but it is not immediately apparent how it can be applied: when does one know that the residual representation is modular?

One example where residual modularity is known is the following: a theorem of Langlands and Tunnel states that any Galois representation (of any number field) into $\mathrm{GL}_2(\mathbf{F}_3)$ is modular. Their result is specific to $\mathbf{F}_3$ and does not apply to representations valued in other finite fields (except perhaps $\mathbf{F}_2$?): the key point is that $\mathrm{GL}_2(\mathbf{F}_3)$ is solvable. Modularity lifting thus allows one to conclude (under appropriate hypotheses) that representations into $\mathrm{GL}_2(\mathbf{Z}_3)$ are modular. Wiles' original application of modularity lifting to elliptic curves used this line of reasoning.

For finite fields other than $\mathbf{F}_3$ (and maybe $\mathbf{F}_2$) there is no analogue of the Langlands–Tunnel theorem: the finite groups $\mathrm{GL}_2(\mathbf{F}_q)$ are typically not solvable. However, Taylor [Tay], [Tay2] partially found a way around this problem: he observed, using a result of Moret-Bailly, that any odd residual representation of a totally real field F becomes modular after passing to a finite extension of F; that is, odd residual representations of F are *potentially modular*. Using modularity lifting, one can conclude that many p-adic are potentially modular as well. Typically, one cannot deduce modularity from potential modularity. Nonetheless, many of the nice properties of modular p-adic representations can be established for potentially modular representations as well: they satisfy the Weil bounds, their L-functions admit meromorphic continuation and satisfy a functional equation, they often can be realized in the Tate module of an abelian variety and they fit into compatible systems. We prove the final of these results.

As if these consequences of potential modularity were not impressive enough, Khare and Wintenberger [KW] went even farther: they proved that every irreducible odd residual representation of $G_{\mathbf{Q}}$ is modular, a result first conjectured by Serre. To do this, they first showed — using potential modularity — that any mod p representation admits a nice p-adic lift. This lift (by one of the corollaries of potential modularity) fits into a compatible system. To prove the modularity of the original mod p representation, it suffices (by modularity lifting, and basic properties of compatible systems) to prove the modularity of the reduction of any of the ℓ-adic representations in the system. This permits the possibility of an inductive argument, which turns out to be quite subtle but possible. The base cases of the induction had been previously proved by Serre and Tate; these results are specific to $\mathbf{Q}$ and is one reason that this sort of result has not been extended to other fields.

Date: March 27, 2010.

As indicated, the results presented here are mainly due to Taylor, Khare and Wintenberger, building on the modularity lifting theorems of Wiles and Kisin (though many other people contributed along the way). I learned most of these arguments by writing a paper [Sno] that extends them a small amount. Some of these notes are taken directly from that paper.

2. Review of compatible systems

In this section we provide a brief review of compatible systems and some of their most basic properties.

2.1. **Compatible systems with rational coefficients.** Let F be a number field. An *n-dimensional compatible system of G_F with coefficients in $\mathbf{Q}$* is a family $\{\rho_\ell\}$ indexed by the set of rational prime numbers ℓ (or possibly some subset thereof) where $\rho_\ell : G_F \to \mathrm{GL}_n(\mathbf{Q}_\ell)$ is a continuous representation, such that the following conditions hold:

- There exists a finite set S of places of F such that each ρ_ℓ is unramified outside $S \cup S_\ell$. Here S_ℓ denotes the set of places of F above ℓ.
- For each place v of F not in S there exists a polynomial $p_v \in \mathbf{Q}[t]$ such that: for any prime ℓ and any place $v \notin S \cup S_\ell$ the characteristic polynomial of $\rho_\ell(\mathrm{Frob}_v)$ is p_v.

In words: the ρ_ℓ have uniform ramification properties and the characteristic polynomial of Frob_v is independent of ℓ.

Example 1. Let f be a Hilbert modular form over F whose Hecke eigenvalues are rational numbers. For each rational prime ℓ we have a Galios representation $\rho_\ell : G_F \to \mathrm{GL}_2(\mathbf{Q}_\ell)$. The collection of these Galois representations forms a compatible system. The set S can be taken to be the set of primes dividing the level of f, while p_v can be taken to be $t^2 - a_v t + a_{v,v}$ where a_v and $a_{v,v}$ are the T_v and $T_{v,v}$ eigenvalues of f.

Example 2. Let E be an elliptic curve over F. Let ρ_ℓ be the representation of G_F on the ℓth Tate module of E (tensored with $\mathbf{Q}_\ell$). Then the collection of these Galois representations forms a two-dimensional compatible system. The set S can be taken to be the set of places of F where E has bad reduction. We have $p_v(t) = t^2 - a_v t + \mathbf{N}\, v$, where $\mathbf{N}\, v + 1 - a_v$ is the number of points of the reduction of E at v with coefficients in the residue field of v. Of course, one can replace E with a higher dimensional abelian variety.

Example 3. Let X be a smooth projective variety over a number field F. Let ρ_ℓ be the representation of G_F on the étale cohomology $H^i(X_{\overline{F}}, \mathbf{Q}_\ell)$, for some fixed i. Then the collection of these Galois representations forms a compatible system. The set S can be taken to be the set of primes where X does not have good reduction. Here, we say that X has good reduction at a place v if there exists a smooth projective scheme $\mathscr{X}/\mathscr{O}_{F_v}$ whose generic fiber is isomorphic to X. The polynomials p_v comes from certain pieces of the zeta function of X (which is by definition independent of ℓ); to find these pieces, the Riemann hypothesis (proved by Deligne) is needed. When X is an abelian variety, this example is more or less the same as the previous one.

Remark 4. Since the compatibility condition is in terms of characteristic polynomials, it is not good at detecting extensions: if $\{\rho_\ell\}$ is a compatible system then so too is $\{\rho_\ell^{\mathrm{ss}}\}$ where ρ^{ss} denotes the semi-simplification of ρ. The converse is not quite true since the ramification of ρ_ℓ cannot be controlled in terms of that of ρ_ℓ^{ss}. We say that a compatible system is *semi-simple* if all of its members are.

2.2. **Compatible systems with general coefficients.** As suggested by the terminology of the previous section, there is a more general notion of compatible system. Let K be a number field. Then an *n-dimensional compatible system of G_F with coefficients in K* is a family $\{\rho_w\}$ indexed by the set of finite places w of K (or possibly some subset thereof) where $\rho_w : G_F \to \mathrm{GL}_n(K_w)$ is a continuous representation, such that conditions analogous to those given in the $K = \mathbf{Q}$ case hold. The polynomial p_v will now have coefficients in K.

Example 5. Let f be a Hilbert modular form over F whose Hecke eigenvalues generate the number field K. Then for each place w of K we have a Galois representation $\rho_w : G_F \to \mathrm{GL}_2(K_w)$, and these form a compatible system. The description of S and p_v are as in Example 1.

Example 6. A *$\mathrm{GL}_2(K)$-type abelian variety* is an abelian variety A/F of dimension $[K : \mathbf{Q}]$ equipped with an injection $\mathscr{O}_K \to \mathrm{End}(A)$. This implies that $T_\ell A \otimes \mathbf{Q}_\ell$ is a free $K \otimes \mathbf{Q}_\ell$ module of rank two. Decomposing this module into its pieces (corresponding to how ℓ splits in K), gives a two dimensional Galois representation $G_F \to \mathrm{GL}_2(K_w)$ for each finite place w of K. These form a compatible system.

2.3. **Properties of compatible systems.** The Chebotarev density theorem immediately yields the following result:

Proposition 7. *Let $\{\rho_w\}$ and $\{\rho'_w\}$ be two semi-simple compatible systems of G_F with coefficients in the same field K. Assume there is some place w_0 of K such that ρ_{w_0} and ρ'_{w_0} are isomorphic. Then ρ_w and ρ'_w are isomorphic for all w.*

As a corollary, we obtain the following, which we will use constantly:

Proposition 8. *Let $\{\rho_w\}$ be a two-dimensional compatible system of semi-simple representations of G_F with coefficients in K, and let w_1 and w_2 be two places of K. Then ρ_{w_1} is modular if and only if ρ_{w_2} is. In particular, if any member of a compatible system is modular then all members are.*

Proof. This follows from the previous proposition since modular representations always come in compatible systems; see Example 5. □

3. Potential modularity

In this section, we sketch give a sketch of Taylor's potential modularity. The original arguments are in the papers [Tay] and [Tay2]. The basic idea is as follows. We are given a two dimensional mod p Galois representation $\overline{\rho}$ of G_F, where F is totally real, which we want to show is potentially modular. We find a two dimensional ℓ-adic Galois representation σ, which we know to modular. This new representation is completely independent of $\overline{\rho}$. However, using a very general theorem of Moret-Bailly, we show that there is a GL_2-type abelian variety A over some finite extension F'/F whose mod ℓ representation is $\overline{\sigma}|_{F'}$ and whose mod p representation is $\overline{\rho}|_{F'}$. Modularity lifting implies that the ℓ-adic representation of A is modular. General properties of compatible systems then give the modularity of the p-adic representation of A, and thus of the mod p representation $\overline{\rho}|_{F'}$ as well.

We now make this precise. We begin by recalling the theorem of Moret-Bailly [MB]:

Theorem 9 (Moret-Bailly)**.** *Let X be a smooth geometrically irreducible variety over a number field F. Let S be a finite set of places of F and for each $v \in S$ let L_v/F_v be a finite Galois extension and let $U_v \subset X(F_v)$ be a non-empty open subset (for the v-adic topology). Then there exists a finite Galois extension F'/F which splits over each L_v (i.e., $F' \otimes_F L_v$ is a direct product of L_v's) and a point $x \in X(F')$ such that the image of x in $X(L_v)$ under any map $F' \to L_v$ belongs to U_v.*

Using this result, we deduce the following crucial result, which "links" arbitrary residual representations.

Proposition 10. *Let F be a totally real number field and let $\overline{\rho}_1 : G_F \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$ and $\overline{\rho}_2 : G_F \to \mathrm{GL}_2(\overline{\mathbf{F}}_\ell)$ be irreducible odd representations, with $p \neq \ell$. Then there exists a finite totally real Galois extension F'/F and a two-dimensional compatible system $\{\rho_w\}$ of representations of $G_{F'}$ with coefficients in some number field K, such that for some place $v_1 \mid p$ of K the representation $\overline{\rho}_{v_1}$ is equivalent to $\overline{\rho}_1$ while for some place $v_2 \mid \ell$ the representation $\overline{\rho}_{v_2}$ is equivalent to $\overline{\rho}_2$. Furthermore, the field F'/F can be taken to be linearly disjoint from any given finite extension of F and the system $\{\rho_w\}$ can be taken so that ρ_{v_1} (resp. ρ_{v_2}) is ordinary crystalline at all places over p (resp. ℓ).*

Proof. For simplicity we assume that $\overline{\rho}_1$ and $\overline{\rho}_2$ take values in $\mathrm{GL}_2(\mathbf{F}_p)$ and $\mathrm{GL}_2(\mathbf{F}_\ell)$ respectively, and that both have cyclotomic determinant. We give some comments on the general case following the proof.

Let Y/F be the moduli space classifying elliptic curves whose p-torsion is $\overline{\rho}_1$ and whose ℓ-torsion is $\overline{\rho}_2$. More precisely, regard $\overline{\rho}_1$ and $\overline{\rho}_2$ as finite étale group schemes G_1 and G_2 over F. Pick an isomorphism $G_1 \to G_1^\vee$ of G_1 with its Cartier dual such that the corresponding pairing $G_1 \times G_1 \to \mathbf{G}_m$ is symplectic, which is possible by the assumption on the determinant of ρ_1; do the same for G_2. For a scheme T/F let $Y(T)$ be the groupoid of elliptic curves E/T equipped with isomorphisms $E[p] \to (G_1)_T$ and $E[\ell] \to (G_2)_T$ such that the Weil pairing on $E[p]$ corresponds to the given pairing on $(G_1)_T$, and similarly for ℓ. It is not difficult to see that Y is representable by a scheme. In fact, the open modular curve $Y(p\ell)$ of full level splits into a several connected components over $\overline{\mathbf{Q}}$ and our space Y is a twisted form of any one of these components. This shows that Y is smooth and geometrically irreducible.

We are now going to apply the theorem of Moret-Bailly. Take S to be the set of infinity places of F and for $v \in S$ let $L_v = F_v$, the real numbers, and let $U_v = Y(L_v)$. Clearly, U_v is an open subset of $Y(L_v)$. To apply the theorem we need $Y(F_v)$ to be non-empty. This is the case because the representations $\overline{\rho}_1$ and $\overline{\rho}_2$ are odd: if E/F_v is any elliptic curve then $E[p]$ is automatically equivalent to $(G_1)_{F_v}$, and similarly for $E[\ell]$.

Thus any elliptic curve over F_v can be given the additional structure needed to define a point of $Y(F_v)$. Moret-Bailly now gives a finite totally real Galois extension F'/F (totally real because it splits over each F_v for $v \mid \infty$) and an elliptic curve E/F' such that $E[p] = \overline{\rho}_1|_{F'}$ and $E[\ell] = \overline{\rho}_2|_{F'}$. The compatible system can now be taken to be the Tate modules of E.

We now show that E may be taken to be ordinary crystalline at all places above ℓ. (The arguments at p are identical, and can be carried out simultaneously.) Add to S all the places of F above ℓ. Fix for the moment a place v of F over ℓ. Let $\overline{U}_v$ be the subset of $Y(\overline{F}_v)$ consisting of elliptic curves with good ordinary reduction. This is clearly a non-empty set, since there exist elliptic curves with good ordinary reduction, and these can be given arbitrary level structure over $\overline{F}_v$. We now show that it is open. Let $j : Y(\overline{F}_v) \to \overline{F}_v$ be the j-invariant; it is a continuous function for the v-adic topology. The subset V of $Y(\overline{F}_v)$ where the elliptic curve has good reduction consists of those curves for which j is integral; it is therefore open. The subset of V where the elliptic curve has ordinary reduction is open, since this only depends upon the reduction of the curve: if E and E' are two curves whose j-invariants are v-adically close then they have the same reduction, and so one is ordinary if and only if the other is. This shows that $\overline{U}_v$ is open. Let L_v/F_v be any Galois extension such that $\overline{U}_v \cap Y(L_v)$ is non-empty, and take U_v to be this intersection. We now apply Moret-Bailly as before. The elliptic curve E/F' that we produce has good ordinary reduction at all places over ℓ by the construction of the sets U_v, and so the Tate module ρ_{v_2} is ordinary crystalline at all places over ℓ.

Finally, we show that F'/F can be taken linearly disjoint from any given finite extension of F. Thus let M/F be a finite extension, which we can and do assume to be Galois. Observe that $Y(F_v)$ is non-empty for all sufficiently large v: indeed, if v is sufficiently large then Y will be smooth at v and its reduction will have rational points by the Weil bounds; smoothness allows us to lift these mod v points to $\mathscr{O}_{F_v}$ points. Let S' be a finite collection of finite places of F satisfying the following conditions: (1) for each $v \in S'$ the set $Y(F_v)$ is non-empty; (2) no place of S' lies over p or ℓ; and (3) no place of S' ramifies in M; (4) the elements Frob_v with $v \in S'$ generate the finite group $\mathrm{Gal}(M/F)$. We now again modify the Moret-Bailly set-up. We add the set S' to the set S, and for $v \in S'$ we take $L_v = F_v$ and $U_v = Y(F_v)$. The field F'/F that Moret-Bailly produces splits at all elements of S' and is therefore linearly disjoint from M. □

Remark 11. In the above proof we assume that $\overline{\rho}_1$ and $\overline{\rho}_2$ had cyclotomic determinant and were valued in the prime field. The first of these conditions is straightforward to relax by passing to an appropriate finite extension of F and twisting. To remove the second assumption one proceeds as follows. Pick a number field K which is sufficiently large so that $\overline{\rho}_1$ can be regarded as taking values in the residue field of K at some place above p, and similarly for ℓ. Then, instead of considering moduli spaces of elliptic curves, consider moduli spaces of $\mathrm{GL}_2(K)$-type abelian varieties. The theory of these moduli spaces is developed in [Rap].

Remark 12. In the previous theorem we required that $\overline{\rho}_1$ and $\overline{\rho}_2$ be irreducible. This is not really needed, but we included since we have defined compatible systems to be rational objects, and so one can typically only form the semi-simplification of their reductions.

We now produce a large supply of "universally" modular Galois representations.

Proposition 13. *Let F be a totally real field and ℓ a prime number. There exists a Galois representation $\sigma : G_F \to \mathrm{GL}_2(\overline{\mathbf{Q}}_\ell)$ satisfying the following conditions:*

(a) *σ is modular.*
(b) *σ is ordinary and crystalline at all places above ℓ.*
(c) *$\overline{\sigma}|_{F(\zeta_\ell)}$ is (absolutely) irreducible.*

Furthermore, these conditions hold after restricting σ to any finite totally real extension of F.

Proof. An exercise in class field theory allows one to produce an imaginary quadratic extension E/F and a character $\psi : G_E \to \overline{\mathbf{Q}}_\ell^\times$ such that the representation $\sigma = \mathrm{Ind}_E^F(\psi)$ satisfies conditions (b) and (c) of the proposition. (One picks E to split at all places of F above ℓ. If $v \mid \ell$ is a place of F and w_1 and w_2 the two places of E above F then one takes ψ so that $\psi|_{E_{w_1}}$ is finitely ramified and $\psi|_{E_{w_2}}$ differs from the cyclotomic character by a finitely ramified character.) A theorem of Hecke states that σ is modular. If F'/F is a finite extension then $\sigma|_{F'} = \mathrm{Ind}_{EF'}^{F'}(\psi|_{F'})$ and the same arguments apply. □

We can now prove potential modularity for residual representations:

Theorem 14. *Let F be a totally real field and let $\overline{\rho} : G_F \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$ be an odd representation such that $\overline{\rho}|_{F(\zeta_p)}$ is (absolutely) irreducible. Then there exists a finite totally real Galois extension F'/F, which can be taken to be linearly disjoint from any finite extension of F, such that $\overline{\rho}|_{F'}$ is modular. Furthermore, the modular form can be taken to be ordinary at all places of F' above p and of level prime to p.*

Proof. Let σ be as in Proposition 13, where ℓ can be any prime different from p (and maybe larger than 5). By Proposition 10, we can find a finite extension F'/F, linearly disjoint from whatever we want, a compatible system $\{\rho_w\}$ of representations of $G_{F'}$ with coefficients in some number field K and two places $v_1 \mid p$ and $v_2 \mid \ell$ of K such that: (1) ρ_{v_1} is ordinary crystalline at all places above p and its reduction is equivalent to $\overline{\rho}$; (2) ρ_{v_2} is ordinary crystalline at all places above ℓ and its reduction is equivalent to $\overline{\sigma}$. The modularity lifting theorem that we have proved now establishes that ρ_{v_2} itself is modular. By compatibility, ρ_{v_1} is modular (see Proposition 8), and thus $\overline{\rho}$ is as well. Since ρ_{v_1} is ordinary crystalline at all places above p, the modular form giving rise to it has prime to p level and is ordinary at all places above p. □

We can now prove potential modularity for p-adic representations:

Theorem 15. *Let F be a totally real field, let $p > 5$ be a prime and let $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ be a continuous representation satisfying the following conditions:*

(A1) *ρ is odd.*
(A2) *ρ ramifies at only finitely many places.*
(A3) *$\overline{\rho}|_{F(\zeta_p)}$ is (absolutely) irreducible.*
(A4) *ρ is ordinary crystalline at all places above p.*

Then there exists a finite totally real Galois extension F'/F, which can be taken to be linearly disjoint from any given finite extension of F, such that $\rho|_{F'}$ is modular.

Proof. By the previous theorem, we can find a finite extension F'/F such that $\overline{\rho}|_{F'}$ comes from a modular form which is ordinary crystalline at all places above p. The modularity lifting theorem we have proved gives the modularity of $\rho|_{F'}$. □

Remark 16. Condition (A4) be relaxed if one is willing to use more general modularity lifting theorems. However, (A1)–(A3) are essential to the method of proof.

Remark 17. This clause about being able to produce the field F' so that it is linearly disjoint from a given extension of F is often used to make F' linearly disjoint from the kernel of $\overline{\rho}$. This implies that $\overline{\rho}$ and $\overline{\rho}|_{F'}$ have the same image. Thus $\overline{\rho}|_{F'}$ will still be irreducible.

4. Putting representations into compatible systems

We now use potential modularity to put p-adic representations in compatible systems. I learned the proof of this result from a lecture given by Taylor at the Summer School on Serre's Conjecture held at Luminy in 2007. Taylor attributed the proof to Dieulefait; a sketch of the argument can be found in [Die, §3.2]. However, I have not found a detailed proof in the literature.

Proposition 18. *Let F be a totally real field, let $p > 5$ be a prime and let $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ be a continuous representation satisfying (A1)–(A4). Then there exists a compatible system $\{\rho_w\}$ of G_F with coefficients in some number field K such that for some place v_0 of K the representation ρ_{v_0} is equivalent to ρ.*

Proof. Apply Theorem 15 to produce a finite Galois totally real extension F'/F linearly disjoint from $\ker \overline{\rho}$ and a modular form f over F' such that $\rho|_{F'} = \rho_f$ (we regard the coefficient field of f as being embedded in $\overline{\mathbf{Q}}_p$). Let I be the set of fields F'' which are intermediate to F' and F and for which $\mathrm{Gal}(F'/F'')$ is solvable. For $i \in I$ we write F_i for the corresponding field. For each i we can use solvable descent to find a modular form f_i such that $\rho|_{F_i} = \rho_{f_i}$. Let K_i denote the field of coefficients of f_i, which we regard as being embedded in $\overline{\mathbf{Q}}_p$. Let K be a number field which is Galois over $\mathbf{Q}$, into which each K_i embeds and which contains all roots of unity of order $[F' : F]$. Fix an embedding $K \to \overline{\mathbf{Q}}_p$ and embeddings $K_i \to K$ such that the composite $K_i \to K \to \overline{\mathbf{Q}}_p$ is the given embedding. Let v_0 be the place of K determined by the embedding $K \to \overline{\mathbf{Q}}_p$. For each place v of K and each $i \in I$ we have a representation $r_{i,v} : G_{F_i} \to \mathrm{GL}_2(K_v)$ associated to the modular form f_i. It is absolutely irreducible. Note that after composing r_{i,v_0} with the embedding $\mathrm{GL}_2(K_{v_0}) \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ we obtain $\rho|_{G_{F_i}}$.

By Brauer's theorem, we can write

$$1 = \sum_{i \in I} n_i \operatorname{Ind}_{\operatorname{Gal}(F'/F_i)}^{\operatorname{Gal}(F'/F)}(\chi_i)$$

where the n_i are integers (possibly negative) and the χ_i are characters of $\operatorname{Gal}(F'/F_i)$ valued in $K^\times$. (Here we use the fact that K contains all roots of unity of order $[F' : F]$.) This equality is taken in the Grothendieck group of representations of $\operatorname{Gal}(F'/F)$ over K. Note that by taking the dimension of each side we find $\sum n_i[F_i : F] = 1$.

Let v be a place of K. For a number field M write $\mathcal{C}_{M,v}$ for the category of semi-simple continuous representations of G_M on finite dimensional K_v-vector spaces. The category $\mathcal{C}_{M,v}$ is a semi-simple abelian category. We let $K(\mathcal{C}_{M,v})$ be its Grothendieck group. It is the free abelian category on the set of irreducible continuous representations of G_M on K_v-vector spaces. We let $(,)$ be the integer valued pairing on $K(\mathcal{C}_{M,v})$ given by $(A, B) = \dim_{K_v} \operatorname{Hom}(A, B)$. This is well-defined because $\mathcal{C}_{M,v}$ is semi-simple. It is symmetric. If M'/M is a finite extension then we have adjoint functors $\operatorname{Ind}_{M'}^{M} : \mathcal{C}_{M',v} \to \mathcal{C}_{M,v}$ and $\operatorname{Res}_{M'}^{M} : \mathcal{C}_{M,v} \to \mathcal{C}_{M',v}$. (One must check, of course, that induction and restriction preserve semi-simplicity — we leave this to the reader.) These functors induce maps on the K-groups which are adjoint with respect to $(,)$. If M_1 and M_2 are two extensions of M and r_1 belongs to $\mathcal{C}_{M_1,v}$ and r_2 belongs to $\mathcal{C}_{M_2,v}$ then we have the formula

$$(\operatorname{Ind}_{M_1}^{M}(r_1), \operatorname{Ind}_{M_2}^{M}(r_2)) = \sum_{g \in S} (\operatorname{Res}_{M_1^g M_2}^{M_1^g}(r_1^g), \operatorname{Res}_{M_1^g M_2}^{M_2}(r_2)) \tag{1}$$

where S is a set of representatives for $G_{M_1} \backslash G_M / G_{M_2}$, M_1^g is the field determined by $gG_{M_1}g^{-1}$ and r_1^g is the representation of $gG_{M_1}g^{-1}$ given by $x \mapsto r_1(g^{-1}xg)$. This formula is gotten by using Frobenius reciprocity and Mackey's formula.

Define

$$\rho_v = \sum_{i \in I} n_i \operatorname{Ind}_{F_i}^{F}(r_{i,v} \otimes \chi_i),$$

which is regarded as an element of $K(\mathcal{C}_{F,v})$. We now show that each ρ_v is (the class of) an absolutely irreducible two dimensional representation. To begin with, we have

$$\begin{aligned}
\rho_{v_0} \otimes_{K_{v_0}} \overline{\mathbf{Q}}_p &= \sum_{i \in I} n_i \operatorname{Ind}_{F_i}^{F}((r_{i,v_0} \otimes_{K_{v_0}} \overline{\mathbf{Q}}_p) \otimes_K \chi_i) \\
&= \sum_{i \in I} n_i \operatorname{Ind}_{F_i}^{F}((\rho|_{F_i}) \otimes_K \chi_i) \\
&= \Big[\sum_{i \in I} n_i \operatorname{Ind}_{F_i}^{F}(\chi_i) \Big] \otimes_K \rho \\
&= \rho
\end{aligned}$$

This shows that ρ_{v_0} is (the class of) an absolutely irreducible representation.

Now let v be an arbitrary finite place of K. We have

$$\begin{aligned}
(\rho_v, \rho_v) &= \sum_{i,j \in I} n_i n_j (\operatorname{Ind}_{F_i}^{F}(r_{i,v} \otimes \chi_i), \operatorname{Ind}_{F_j}^{F}(r_{j,v} \otimes \chi_j)) \\
&= \sum_{i,j \in I} \sum_{g \in S_{ij}} n_i n_j (\operatorname{Res}_{F_i^g F_j}^{F_i^g}((r_{i,v} \otimes \chi_i)^g), \operatorname{Res}_{F_i^g F_j}^{F_j}(r_{j,v} \otimes \chi_j))
\end{aligned}$$

where we have used (1). Here S_{ij} is a set of representatives for $G_{F_1} \backslash G_F / G_{F_2}$. The representation $r_{i,v}|_{F'}$ is the representation coming from the form f' and so is absolutely irreducible. It follows that the restriction of $r_{i,v}$ to any subfield of F' is absolutely irreducible. Thus the representations occurring in the pairing in the second line above are irreducible. It follows that the pairing is then either 1 or 0 if the representations are isomorphic or not. Therefore, if let $\delta_{v,i,j,g}$ be 1 or 0 according to whether $\operatorname{Res}_{F_i^g F_2}^{F_i^g}(r_{i,v} \otimes \chi_i)^g$ is isomorphic to $\operatorname{Res}_{F_i^g F_2}^{F_j}(r_{j,v} \otimes \chi_j)$ then we find

$$(\rho_v, \rho_v) = \sum_{i,j \in I} \sum_{g \in S_{ij}} n_i n_j \delta_{v,i,j,g}.$$

Now, the $\{r_{i,v}\}_v$ and the $\{r_{j,v}\}_v$ form a compatible system. It follows that $\delta_{v,i,j,g}$ is independent of v. The above formula thus gives

$$(\rho_v, \rho_v) = (\rho_{v'}, \rho_{v'})$$

if v' is another place of K. Taking $v' = v_0$ and using that ρ_{v_0} is an absolutely irreducible representation gives $(\rho_v, \rho_v) = 1$. Now, if we write $\rho_v = \sum m_i \pi_i$ where $m_i \in \mathbf{Z}$ and the π_i are mutually non-isomorphic irreducible representations then we have $(\rho_v, \rho_v) = \sum m_i^2 (\pi_i, \pi_i)$. Since the terms are all non-negative integers and the sum is 1, we find $\rho_v = \pm\pi$ with $(\pi, \pi) = 1$. Thus π is an absolutely irreducible representation. Now, $\dim \rho_v = 2$ since each $r_{i,v}$ is two dimensional and $\sum n_i [F_i : F] = 1$. Since $\dim \pi$ is non-negative, we must have $\rho_v = \pi$. This proves that ρ_v is the class of an absolutely irreducible representation.

Of course, it must be shown that the ρ_v actually form a compatible system! This is fairly easy after what we have done, and we leave this task to the reader. □

Remark 19. The compatible system constructed above is in fact *strongly compatible*. For a discussion of this, see [Tay, Theorem 6.6].

5. Lifting residual representations

We now show that one can lift most residual representations to characteristic zero representations.

Proposition 20. *Let F be a totally real field, $p > 5$ a prime and $\overline{\rho} : G_F \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$ an odd representation such that $\overline{\rho}|_{F(\zeta_p)}$ is (absolutely) irreducible. Assume that for each plave $v \mid p$ of F the representation $\overline{\rho}|_{F_v}$ admits a lift to $\overline{\mathbf{Z}}_p$ which is ordinary crystalline. Then there exists a continuous representation $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ satisfying (A1)–(A4) lifting $\overline{\rho}$. One can take ρ to be unramified at the same places where $\overline{\rho}$ is unramified (excluding places above p).*

Proof. Let S be the set of primes away from p at which $\overline{\rho}$ ramifies and let S_p denote the set of primes above p. For $v \in S \cup S_p$ we have the universal framed deformation ring $R_v^{\square}$ of $\overline{\rho}|_{F_v}$. For $v \in S_p$ we let $R_v^{\dagger}$ be the quotient of $R_v^{\square}$ parameterizing ordinary crystalline representations, in the same manner as we have done before. The ring $R_v^{\dagger}$ is non-zero since we have assumed that $\overline{\rho}|_{F_v}$ admits an ordinary crystalline lift. Our previous work therefore shows that it is $\mathscr{O}$-flat and has relative dimension dimension is $3 + [F_v : \mathbf{Q}_p]$ over $\mathscr{O}$. For $v \in S$ we pick a non-zero $\mathscr{O}$-flat quotient $R_v^{\dagger}$ of $R_v^{\square}$ of relative dimension 3 over $\mathscr{O}$. It takes a little bit of work to show that such a quotient exists, but it is not very hard. (The calculations appear in [Sno], and they probably are also somewhere in [KW].) We let $\widetilde{B}$ (resp. B) be the completed tensor product of the $R_v^{\square}$ (resp. $R_v^{\dagger}$) for $v \in S \cup S_p$. We let $R^{\square}$ be the universal framed deformation ring for $\overline{\rho}$ unramified outside of S. We put $R^{\dagger} = R^{\square} \otimes_{\widetilde{B}} B$ and let $R^{\ddagger}$ be the unframed version of $R^{\dagger}$.

Now, we have a presentation for $R^{\square}$ over $\widetilde{B}$ [Ki, Prop. 4.1.5]:

$$R^{\square} = \widetilde{B}[\![x_1, \dots, x_{r+n-1}]\!]/(f_1, \dots, f_{r+s})$$

where $s = \sum_{v|\infty} \dim H^0(F_v, \mathrm{ad}^\circ \overline{\rho})$, n is the cardinality of $S \cup S_p$ and r is some non-negative integer. Tensoring this over $\widetilde{B}$ with B gives

$$R^{\dagger} = B[\![x_1, \dots, x_{r+n-1}]\!]/(f_1, \dots, f_{r+s})$$

Now, since $\overline{\rho}$ is odd, we have $s = [F : \mathbf{Q}]$. On the other hand, the dimension of B is $[F : \mathbf{Q}] + 3n + 1$. We conclude that $R^{\dagger}$ has dimension at least $4n$. Since $R^{\dagger}$ is a power series ring over $R^{\ddagger}$ in $4n - 1$ variables, we find that $R^{\ddagger}$ has dimension at least 1.

Let F'/F be a finite totally real extension over which $\overline{\rho}$ becomes modular, by an ordinary modular form of level prime to p. We can then define deformation rings for $\overline{\rho}|_{F'}$ analgous to the ones we have defined for $\overline{\rho}$. We will denote these rings with an overline. There is a natural map $\overline{R} \to R$ (the universal unframed deformation rings unramified outside of S), which is easily verified to be a finite map of rings. It follows that the induced map $\overline{R}^{\ddagger} \to R^{\ddagger}$ is finite as well. Now, to establish our modularity lifting theorem we identified $\overline{R}^{\ddagger}[1/p]$ with a Hecke algebra using a patching argument. Out of this argument we obtained another piece of information: that $\overline{R}^{\ddagger}$ itself, without p inverted, is finite over $\mathscr{O}$. (Actually, we did not quite use the ring $\overline{R}^{\ddagger}$, we needed to make a slight modification of the local deformation ring at p. Nonetheless, the same argument establishes the finiteness of $\overline{R}^{\ddagger}$.) We now conclude that $R^{\ddagger}$ itself is finite over $\mathscr{O}$.

We have thus show that $R^\ddagger$ is finite over $\mathscr{O}$ and has Krull dimension at least one. These two properties imply that $R^\ddagger$ cannot consist soley of p-power torsion. Therefore $R^\ddagger[1/p]$ is non-zero, and so there exists some homomorphism $R^\ddagger \to \overline{\mathbf{Q}}_p$. The corresponding deformation is the representation ρ that we are required to produce. □

Remark 21. One can still prove that $\overline{\rho}$ admits a nice lift without the assumption that $\overline{\rho}|_{F_v}$ admit an ordinary crystalline lift for each $v \mid p$. Of course, without this assumption the resulting lift cannot be assured to be ordinary. Not surprisingly, this more general statement makes use of more general modularity lifting theorems.

6. Remarks on Serre's conjecture

Recall Serre's conjecture:

Conjecture 22. *Any odd semi-simple representation $\overline{\rho} : G_{\mathbf{Q}} \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$ is modular.*

When $\overline{\rho}$ is reducible it is easy to see that it is easy to see that $\overline{\rho}$ is modular. This is very far from the case when $\overline{\rho}$ is irreducible. However, Khare and Wintenberger proved this a few years ago. We now give some idea of the proof.

To begin with, Serre made a stronger conjecture, specifying the optimal weight and level of a modular form giving rise to $\overline{\rho}$. (See Akshay's talk for more details along these lines.) The level $N(\overline{\rho})$ is just the prime-to-p Artin conductor of $\overline{\rho}$. Thus is $N(\overline{\rho})$ always prime to p, and $\ell \mid N(\overline{\rho})$ if and only if $\overline{\rho}$ is ramified at ℓ. The weight $k(\overline{\rho})$ is more complicated to define, but it can be bounded in terms of p. It is known that if $\overline{\rho}$ is modular then it is modular of this optimal weight and level. Furthermore, the work we have done in §4 and §5 can be generalized to show that $\overline{\rho}$ lifts to a strongly compatible system of weight $k(\overline{\rho})$ and conductor $N(\overline{\rho})$. This uses more advanced modularity lifting theorems. (The weight of a p-adic representation is defined using p-adic Hodge theory. The conductor of a p-adic representation is a product of the usual prime-to-p part together with a p-part coming from p-adic Hodge theory. If $\{\rho_\ell\}$ is a strongly compatible system then all the ρ_ℓ have the same weight and conductor.)

We begin by discussing the level one case of Serre's conjecture. We have the following result:

Proposition 23 (Serre, Tate)**.** *Conjecture 22 holds if $N(\overline{\rho}) = 1$ and $p = 2$ or $p = 3$.*

The $p = 2$ case is due to Tate, the $p = 3$ case to Serre. In fact, there are no cusp forms of level 1 and small weight, so the above proposition is really saying that there are no irreducible representations $G_{\mathbf{Q}} \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$ ramified only at p for $p = 2, 3$.

This result allows one to try to attempt an inductive argument. Let $\overline{\rho} : G_{\mathbf{Q}} \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$ have $N(\overline{\rho}) = 1$. Lift $\overline{\rho}$ to a compatible system $\{\rho_\ell\}$ of conductor 1 and weight $k = k(\overline{\rho})$. By the above result, we know that the reduction of ρ_3 is modular. We would like to use a modularity lifting theorem to conclude that ρ_3 is modular. Of course, this is going to require a more powerful modularity lifting theorem than we have discussed. Such theorems do exist (and can handle, for instance, the fact that $\overline{\rho}_3$ will be reducible), but they are not completely unconditional: the weight has to be small compared to p. Thus if one is going to apply a modularity lifting theorem in characteristic 3 the weight has to be quite small (maybe 3 or 4). Our compatible system $\{\rho_\ell\}$ can have arbitrarily large weight, so this is a real problem! (One might think to try to lift our original $\overline{\rho}$ to a small weight p-adic representation unramified outside of p, and then put this in a compatible system. This is possible, but the small weight p-adic representation will typically have a conductor at p; this means that the 3-adic representation will ramify at p and we can no longer use the theorems of Serre and Tate.)

To get around this problem, Khare (who proved the level one case before he and Wintenberger established the general case) employs an inductive argument on the weight and the prime. I do not know the details of how this works, so I cannot explain it.

Now consider the general case, where $N(\overline{\rho})$ is no longer assumed to be 1. The proof of Khare–Wintenberger is again an induction, but now the level is considered as well. Here is one way the level can be cut down: lift $\overline{\rho}$ to a compatible system $\{\rho_\ell\}$. Say $\ell \mid N(\overline{\rho})$. Then look at $\overline{\rho}_\ell$. By definition, its Serre-level is prime to ℓ. If we were just inducting on the level, then we could assume that $\overline{\rho}_\ell$ were modular (since it has smaller level than $\overline{\rho}$). Of course, we would then like to conclude that ρ_ℓ is modular as well. However, the available modularity lifting theorems may not be strong enough for us to make this deduction — for instance, the weight could be too larger compared to ℓ. I think the argument of Khare–Wintenberger runs induction on

several things at once to get around this sort of issue. Again, I do not know the details, so I will leave it at that.

References

[Die] L. Dieulefait, *Uniform behavior of families of Galois representations on Siegel modular forms and the endoscopy conjecture*, Bol. Soc. Mat. Mexicana (3) **13** (2007), no. 2, 243–253.

[Ki] M. Kisin, *Modularity of 2-dimensional Galois representations*, Current Developements in Mathematics (2005), 191–230.

[KW] C. Khare and J.-P. Wintenberger, *Serre's modularity conjecture (II)*, preprint, 76 pages.

[MB] L. Moret-Bailly, *Groupes de Picard et problèmes de Skolem II.*, Ann. Sci. de l'É.N.S. (4) **22** (1989) no. 2, 181–194.

[Rap] M. Rapoport, *Compactifications de l'espace de modules de Hilbert-Blumental*, Compositio Math. **36** (1978), no. 3, 255–335.

[Tay] R. Taylor, *Remarks on a conjecture of Fontaine and Mazur*, J. Inst. Math. Jussieu 1 (2002), no. 1, 125–143.

[Tay2] R. Taylor, *On the meromorphic continuation of degree two L-functions*, Documenta Math., Exta Volume: John H. Coates' Sixtieth Birthday (2006), 729–779 (electronic).

[Sno] A. Snowden, *Weight two representations of totally real fiels*, arxiv/0905.4266.

The Patching Argument

Brandon Levin

August 2, 2010

1 Motivating the Patching Argument

My main references for this talk were Andrew's overview notes and Kisin's paper "Moduli of Finite Flat Group Schemes." I also would like to thank Andrew for his help and letting me incorporate some of his Tex code which saved me time and energy.

Since this is the final lecture in this seminar, we begin by stating the theorem we set out to prove.

Theorem 1.1. *Let $F/\mathbb{Q}$ be a totally real number field and let $\rho : G_F \to \mathrm{GL}_2(\bar{\mathbb{Q}}_p)$ be a continuous representation of its absolute Galois group, with $p > 5$. Assume that ρ satisfies the following conditions:*

- *ρ ramifies at only finitely many places.*
- *ρ is odd, i.e., $\det \rho(c) = -1$ for all complex conjugations $c \in G_F$.*
- *ρ is potentially crystalline and ordinary at all places above p.*
- *$\bar{\rho}|_{G_{F(\zeta_p)}}$ is absolutely irreducible.*
- *There exists a parallel weight two Hilbert modular form f such that ρ_f is potentially crystalline and ordinary at all places above p and $\bar{\rho} = \bar{\rho}_f$.*

Then there exists a Hilbert modular form g such that $\rho = \rho_g$.

As explained in Andrew's Lecture 18, we can make a solvable totally real base changes to arrange so that ρ is crystalline and ordinary at all places dividing p and that ρ is Steinberg at all places where it is ramified. We are careful to choose a solvable extension preserving the absolute irreducibility condition. Solvable base change ensures that modularity of this new ρ implies modularity of the one we started with. All the other conditions are preserved. For convenience, we also base change so that the number of real places of F and the number of Steinberg places of ρ are both even.

The techniques of level lowering and level raising as discussed in Akshay's most recent talk allows us find a new Hilbert modular form f' with the same reduction mod p whose level exactly matches with ρ, that is, f' is ramified only where ρ is Steinberg and is Steinberg there, and further f' is ordinary at all places dividing p. Some further reductions are discussed in Andrew's notes to get ourselves to the following situation:

Theorem 1.2. *We have a representation*

$$\bar{\rho} : G_F \to \mathrm{GL}_2(k)$$

where k is a finite field of characteristic p, a finite set St of places of F away from p and a modular representation ρ_f lifting $\bar{\rho}$. Let S_p denote the places of F above p. We assume the following hypotheses:

(A1) *ρ_f is crystalline and ordinary at all places in S_p, Steinberg at all places in St and unramified at all other places.*

(A2) *$\det \rho_f = \chi_p$.*

(A3) *$\bar{\rho}|_{G_{F(\zeta_p)}}$ is absolutely irreducible.*

(A4) *$\bar{\rho}|_{G_{F_v}}$ is trivial for $v \in S_p \cup St$.*

(A5) *F has even degree over $\mathbb{Q}$ and St has even cardinailty.*

Then ρ is modular.

As we go through the patching argument, the question may arise: why would one think to do it this way? The argument feels very unnatural. The best I can do is to point to similar argument from Iwasawa theory which arose much more naturally and undoubtably inspired this one. If you are unfamiliar with Iwasawa theory or not interested, you may skip to the next section.

Everything I say can be found in Washington's book *Cyclotomic Fields* in much more detail. The important first case in Iwasawa theory is the study of the p-part of the class group of the cyclotomic fields $\mathbb{Q}(\zeta_{p^n}) = K_n$. Let M_n be the p-part of the class group K_n. One first observes that M_n comes with an action of $\mathrm{Gal}(K_n/\mathbb{Q})$. For our brief discussion, the relevant action is that of $\mathrm{Gal}(K_n/K_1)$ which is a cyclic group of order p^{n-1}. Furthermore, we think of $\mathrm{Gal}(K_n/K_1)$ as a quotient of $\Gamma = \mathrm{Gal}(K_\infty/K_1)$. So that Γ acts on all the M_n's.

Now, there exists maps of abelian groups

$$M_{n+1} \to M_n.$$

One can think interpret this map as norm map K_{n+1}/K_n either at the level of ideal class groups or at the level of ideles. Or an alternative description exists in terms of Hilbert class fields. It is not hard to show the map is both surjective and Γ-equivariant.

Each M_n is an abelian p-group with the structure of a $\mathbb{Z}_p[\Gamma/\Gamma^{p^{n-1}}]$-module. Hence the projective limit

$$M_\infty := \lim_{\leftarrow} M_n$$

is naturally a module over the completed group ring

$$\lim_{\leftarrow} \mathbb{Z}_p[\Gamma/\Gamma^{p^{n-1}}] = \mathbb{Z}_p[[\Gamma]] \cong \mathbb{Z}_p[[T]]$$

the last isomorphism being given by sending a topological generator for Γ to $1+T$.

The value of this limiting process is that while the M_n themselves may be very mysterious, there is a nice structure theory for $\mathbb{Z}_p[[T]]$-modules which can be applied to M_∞. Knowing this and the fact that M_n can be recovered as quotient of M_∞ by an augmentation ideal, allows one to derive strong results about how $|M_n|$ grows as we go up the tower.

Idea of the Proof: As both Mike and Sam have discussed, given a set of Taylor-Wiles primes Q, the corresponding deformation ring and space of modular forms have an action of $\mathcal{O}[\Delta_Q]$. The naive idea would be to take a limit where the Taylor-Wiles sets become congruent to 1 modulo higher and higher powers of p. The limiting "deformation ring" and "space of forms" then become modules over a power series in $|Q|$ variables. One could then use commutative algebras results which only work over domains to deduce that $R = T$.

One big problem not present in the Iwasawa setup is that these TW-sets have absolutely no relationship to each other. There will be no obvious maps from M_{n+1} to M_n, and this will lead to the "miracle" that is patching argument. While the Iwasawa construction moves vertically, the Taylor-Wiles construction is "horizontal" one and will involve many more choices, but nevertheless, in the end, it works and we can recover facts about our original R and T in exactly the same way one does in Iwasawa theory.

2 Setup/Recollection of Previous Results

There is unfortunately a huge amount of notation to set up here so let's get started.

2.1 Deformation Rings

All our deformation rings will be algebras over $\mathcal{O}$ which is the ring of integers of its fraction field E, a p-adic field. We will denote the residue field of $\mathcal{O}$ by $\mathbb{F}$ and assume the $\bar{\rho}$ is representation over $\mathbb{F}$.

In this section, we add in the framings we need and recall the necessary dimension formulas.

Let $\tilde{R}^{\square}$ be the universal deformation ring of $\bar{\rho}$ unramified outside $S_p \cup St$ together with framings at each place in $S_p \cup St$. The essential thing is that $\tilde{R}^{\square}$ is naturally an algebra

over the universal local framed deformation ring $\tilde{R}_v$ of $\bar{\rho}|_{G_{F_v}}$ for all $v \in S_p \cup St$. All our local deformation rings will be framed so I leave off the box.

The functor of forgetting the framings makes $\tilde{R}^{\square}$ an algebra over the plain old global deformation ring $\tilde{R}$ which exists since $\bar{\rho}$ is absolutely irreducible.

We now state the first important dimension formula which was discussed in the fall:

(R1) $\tilde{R}^{\square}$ is smooth over $\tilde{R}$ of relative dimension $j = 4|S_p \cup St| - 1$.

Note that we need these framings at $v \in S_p \cup St$ because otherwise the local deformation rings may not be representable. In fact, as Andrew points out in his overview we can actually assume that $\bar{\rho}|_{G_{F_v}}$ is trivial.

Now let R_v be the quotient of $\tilde{R}_v$ which represents the crystalline-ordinary (respectively Steinberg component) of the local deformation ring at $v \in S_p \cup St$ as discussed in Rebecca and Brian's (Corollary 4.3) talks earlier this quarter.

Set $\tilde{B} = \hat{\otimes}_{v \in St \cup S_p} \tilde{R}_v$ and $B = \hat{\otimes}_{v \in St \cup S_p} R_v$, and note that B is a quotient of $\tilde{B}$.

The relevant properties of the R_v are

- R_v is a flat $\mathcal{O}$-algebra (also complete local with residue field $\mathbb{F}$).
- R_v is a domain (i.e. $\operatorname{Spec} R_v$ is connected).
- $R_v[1/p]$ is a regular E-algebra.
- R_v has relative dimension 3 for $v \in St$ and $3 + [F_v : Q_p]$ for $v \in S_p$ over $\mathcal{O}$.

These were all discussed or proved in previous lectures except I believe the dimension formula which was only discussed for $R_v[1/p]$ but should be in the notes. For later applications, we will also want to know that R_v has the same dimension at all maximal ideals. This is Lemma 4.6 in Brian's notes from this quarter.

Proposition 2.1. *(B1) B is a flat $\mathcal{O}$-algebra (also complete local with residue field $\mathbb{F}$).*

(B2) B is a domain of relative dimension $3|S_p \cup St| + [F : Q]$ over $\mathcal{O}$.

(B3) $B[1/p]$ is a regular E-algebra.

Proof. Since we can build B up one step at a time, for simplicity, let R_1 and R_2 complete local $\mathcal{O}$-algebras that are domains and which become formally smooth after inverting p. Because the formation of the local deformation rings commutes with any finite extension $\mathcal{O}'$ of $\mathcal{O}$, we can further assume that same properties hold for $R_i \otimes_{\mathcal{O}} \mathcal{O}'$.

Let $A = R_1 \hat{\otimes} R_2$. For (B1), it suffices to show that A if flat over R_1 since R_1 is flat over $\mathcal{O}$. Note that $R_1 \to A$ is a local map of complete local Noetherian rings. By Prop 5.1, it suffices to show that $A/m_1^n A$ is flat over R_1/m_1^n where m_1 is the maximal ideal of R_1. However, once you quotient by m_1^n the completed tensor product goes away and we get

$$A/m_1^n A \cong R_1/m_1^n \otimes_{\mathcal{O}} R_2$$

which is clearly flat over R_1/m_1^n since R_2 is flat over $\mathcal{O}$.

Since A is $\mathcal{O}$-flat, it is p-torsion free and so $A \hookrightarrow A[1/p]$. Thus for (B2) and (B3), it suffices to show $A[1/p]$ is a regular domain. Intuitively, one might think of $R_i[1/p]$ as being bounded functions on the corresponding rigid analytic space which we will call X_i. This is not quite true, but in any case, if it were, then $A[1/p]$ would be bounded functions on the rigid analytic product space $X_1 \times X_2$. And we are reduced to the statement that the product of smooth spaces is smooth and product of geometrically connected is geometrically connected. This motivates why one might believe modulo lots of technical details that it would be true. Now I give the more hands-on algebraic proof.

Let $X_1 = \operatorname{Spec} R_1[1/p], X_2 = \operatorname{Spec} R_2[1/p]$, and $X = \operatorname{Spec} A[1/p]$. The rough idea is that while X is very far from being the product of X_1 and X_2, it looks like a product at the level of MaxSpec and this turns out to be enough. This is made precise in Lemma 5.2.

Recall also the following essential facts from Brian's lecture "Generic Fibers of Deformation Rings":

1. $A[1/p], R_1[1/p], R_2[1/p]$ are all Noetherian and Jacobsen rings. In particular, their closed points are dense in their spectrum.

2. All maximal ideal of $A[1/p], R_1[1/p], R_2[1/p]$ have a residue field a finite extension of E.

3. Under any homomorphism $R_i[1/p] \to E'$ where E' is finite extension of E, R_i lands in the ring of integers of E'.

Regularity (B3) for Noetherian ring over a field can be checked by a functorial criterion on Artin local E-algebras (see Remark 2.3 in Lecture 21). Thus, it follows immediately from Lemma 5.2. Knowing regularity, we get the dimension count by applying Lemma 5.2 to the dual numbers over E. It remains to show the X is connected.

There are natural projections $\pi_i : X \to X_i$ given by the evident ring inclusions. Further, given any rational point $x \in X_2(E)$, we get a section $s_x : X_1 \to X$. Our point x corresponds to a map $R_2[1/p] \to E$ which from fact (3), gives rise to a map

$$R_2 \to \mathcal{O}$$

which induces a map

$$R_1 \hat{\otimes}_{\mathcal{O}} R_2 \to R_1$$

which after inverting p yields s_x. Note that by construction s_x is a closed immersion onto the fiber $\pi_2^{-1}(x)$.

We can now show X is irreducible. Assume that U and V were two disjoint non-empty open subsets of X. After extending the field if necessary, we can assume U and V contain rational points u and v. Let s_u and s_v be sections passing through u and v respectively. Since X_1 is irreducible, $s_u^{-1}(U) \cap s_v^{-1}(V)$ is non-empty. Again extending the field, we can assume it contains a rational point y. Consider the fiber $X_y = \pi_1^{-1}(y)$. Both U and V intersect X_y, a closed subset of X. By remark above, X_y is the image of some section s_y and hence irreducible because X_2 is. $\square$

The universal deformation ring for $\bar{\rho}$ unramified outside $S_p \cup St$, crystalline and ordinary in S_p, and Steinberg in St with local framings is given by

$$R^{\square} = \tilde{R}^{\square} \hat{\otimes}_{\tilde{B}} B.$$

Intuitively, one should just think the B is gotten from $\tilde{B}$ by the univeral equations forcing the desired local properties and all that we are doing here is just applying those universal conditions to the global deformation ring.

Proposition 2.2. *Let*

$$g = \dim(\ker(H^1(G_{F_{S_p \cup St}}, \mathrm{ad}^0) \to \oplus_{v \in S_p \cup St} H^1(G_{F_v}, \mathrm{ad}^0)) + \sum_{v \in S_p \cup St} \dim H^0(F_v, \mathrm{ad}) - \dim H^0(G_{F_{S_p \cup St}}, \mathrm{ad}) \tag{1}$$

Then, $R^{\square}$ can be written as a quotient of $B[[x_1, \ldots, x_g]]$.

Proof. The quantity g is exactly the number of generators of $\tilde{R}$ over $\tilde{B}$. This can be shown by a slight modification (to take into account framings) of the argument given by Samit in Lecture 6. Taking any presentation for $\tilde{R}$ over $\tilde{B}$ and then tensoring with B over $\tilde{B}$ gives the desired presentation. □

2.2 Taylor-Wiles Sets

Just as in Mike's lecture, a *TW set of primes* is a set Q of places of F satisfying the following conditions:

- Q is disjoint from S_p and St.
- $N(v) \equiv 1 \pmod{p}$ for all $v \in Q$.
- The eigenvalues of $\bar{\rho}(\mathrm{Frob}_v)$ are distinct and belong to k.
- The map
$$H^1(G_{F_{S_p \cup St \cup Q}}, \mathrm{ad}^0(\bar{\rho})(1)) \to \oplus_{v \in Q} H^1(F_v, \mathrm{ad}^0(\bar{\rho})(1))$$
is an isomorphism.

Note that the last conditions implies that TW-sets always have the same size h.

Given a TW-set Q, we define $R^{\square}_Q$ to be the universal deformations ring unramified outside $S_p \cup St \cup Q$ which is ord-cryst at S_p and Steinberg at St and with local framings at $S_p \cup St$ but not at Q. So $R^{\square}_Q$ is exactly the same as $R^{\square}$ except we allow ramification now at the auxiliary set Q. Note that there exists a natural map $\varphi_Q : R^{\square}_Q \to R^{\square}$ since $R^{\square}$ is the unramified at Q quotient of $R^{\square}_Q$. We now recall a series of important properties of these $R^{\square}_Q$.

(Q1) All the conditions together combined with a duality result for Selmer groups imply $R^{\square}_Q$ is a quotient of $B[[x_1, \ldots, x_g]]$ just as $R^{\square}$ is.

(Q2) $R_Q^{\square}$ is an algebra over the group ring $\mathcal{O}[\Delta_Q]$, where Δ_Q is the maximal pro-p quotient of $\prod_{v \in Q} \mathcal{O}_v$.

(Q3) The kernel of φ_Q is the augmentation ideal $\mathfrak{a}_Q R_Q^{\square}$ where $\mathfrak{a}_Q$ is the augmentation ideal of $\mathcal{O}[\Delta_Q]$.

Let me say a word or two about these results. (Q2) involves a choice of root of the characteristic polynomial of Frob_v for each $v \in Q$ or equivalently a choice of one of the two univeral tame characters mapping I_v into $R_Q^{\square,\times}$. We will resolve this ambiguity by including the choice in our data below. Since the action of $\mathcal{O}[\Delta_Q]$ is exactly the "univeral" action of inertia, the unramified at Q quotient is given by the co-invariants under the Δ_Q action, that is,

$$R^{\square} \cong R_Q^{\square} / < g - 1 > R_Q$$

where $< g - 1 >$ is ideal generated by running over all $g \in \Delta_Q$. This gives (Q3).

Definition 2.3. A *TW-datum of depth* n is a TW-set Q_n together with a choice of root α_v of the characterstic polynomial of $\bar{\rho}(\mathrm{Frob}_v)$ for each $v \in Q_n$ and such that for all $v \in Q_n$, $Nv \equiv 1 \mod p^n$.

Mike showed in Lecture 24 the existence of TW-datum for any depth n given our assumptions on $\bar{\rho}$. For each $n \geq 1$, we fix once and for all a TW-datum Q_n of depth n. The end result being the existence of $R_{Q_n}^{\square}$ together with an $\mathcal{O}[\Delta_{Q_n}]$ structure. The condition on the norm of $v \in Q_n$ forces Δ_{Q_n} to grow with n.

Given (Q1), we now fix a surjection

$$B[[x_1, \ldots, x_g]] \to R_{Q_n}^{\square}$$

for all Q_n.

$R_{Q_n}^{\square}$ is an algebra over $\mathcal{O}[\Delta_{Q_n}]$ but choosing framing variables (i.e $R_{Q_n}[[y_1, \ldots, y_j]] \cong R_{Q_n}^{\square}$), we can make $R_{Q_n}^{\square}$ an algebra over $\mathcal{O}[[y_1, \ldots, y_j]][\Delta_{Q_n}]$. Choosing generators for the cyclic factors of Δ_{Q_n}, we get a homomorphism

$$\gamma_n : \mathcal{O}[[y_1, \ldots, y_j, T_1, \ldots, T_h]] \to R_{Q_n}^{\square}$$

where the kernel of γ_n contains $(T_i + 1)^{p^{n_i}} - 1$ for some $n_i \geq n$.

The following key formula says that $B[[x_1, \ldots, x_g]]$ and $\mathcal{O}[[y_1, \ldots, y_j, T_1, \ldots, T_h]]$ have the same dimension.

Proposition 2.4. *Let g, h, j be defined as above with g generating global over local, h being size of TW-set, and j generating framed global over global. Then,*

$$h + j + 1 = \dim B + g.$$

Proof. In the end, we will only need an inequality ($\geq$). However, in this case, we know equality and so we might as well prove it. We have

$$h + j - g = h + (4|S_p \cup St| - 1) - (h - [F : \mathbb{Q}] + |S_p \cup St| - 1)$$

by (R1) and Proposition 1. Simplifying we get

$$h + j - g = [F : \mathbb{Q}] + 3|S_p \cup St| = \dim B - 1.$$

□

2.3 Hecke Modules

We now turn to the modular forms side. Let D be the unique quaternion algebra over F ramifying exactly at all infinite places and all places in St. Jacquet-Langlands tells us that any modular form f which could satisfy $\rho = \rho_f$ would come from a form on D. Thus, we lose nothing by working on D where certain things are much simpler.

Recall the following key result which Sam discussed last week:

Proposition 2.5. *Given a TW-datum Q of any depth, there exists a level U and direct summand M_Q of the space of automorphic forms $S(U)$ on D which is a Hecke-stable submodule such that the deformation ring $\tilde{R}_Q$ acts via a natural map $\tilde{R}_Q \to T_Q$. $\tilde{R}_Q$ comes with a $\mathcal{O}[\Delta_Q]$ structure which induces an action on M_Q. M_Q is a finite free $\mathcal{O}[\Delta_Q]$-module. Furthermore, $M = M_Q/\mathfrak{a}M_Q$.*

Let M_{Q_n} be the space of modular forms associated to our TW-datum Q_n.

It is a small technical point, but we have to pass to a framed version of M_{Q_n} to make the argument work. We use our chosen presentation

$$\tilde{R}_{Q_n}[[y_1, \ldots, y_j]] \twoheadrightarrow \tilde{R}^{\square}_{Q_n}$$

which is a map of $\mathcal{O}[\Delta_Q]$-algebras. Define

$$M^{\square}_Q := M_Q \otimes_{\tilde{R}_{Q_n}} \tilde{R}^{\square}_{Q_n}$$

and similarly for M.

Proposition 2.6 (H1)**.** *Using the $\mathcal{O}[[y_1, \ldots, y_j]]$-algebra structure coming from the framing, $M^{\square}_{Q_n}$ is a finite free over $\mathcal{O}[[y_1, \ldots, y_j]][\Delta_{Q_n}]$ and $M^{\square} = M^{\square}_{Q_n}/\mathfrak{a}M^{\square}_{Q_n}$.*

Proof. Consider the following diagram:

$$\begin{array}{ccc} \mathcal{O}[[y_1, \ldots, y_j]][\Delta_{Q_n}] & \longrightarrow & \tilde{R}^{\square}_{Q_n} \\ \uparrow & & \uparrow \\ \mathcal{O}[\Delta_{Q_n}] & \longrightarrow & \tilde{R}_{Q_n} \end{array}$$

which one can show is a Cartesian square by considering the chosen presentation of the unframed over the framed. Further, the vertical arrows are faithfully flat. The first statement follows from flat base change; the second from the fact that quotients commute with flat extension. □

3 Passing to the Limit

We recall now where we are headed.

A priori, we only started with a surjective map of $\mathcal{O}$-algebras

$$\varphi : \tilde{R} \to T$$

where T is some Hecke algebra acting faithfully on M a space of modular forms. $\tilde{R}$ is the deformation ring with no local conditions. We can pass to the framed version of this map:

$$\varphi^{\square} : \tilde{R}^{\square} \to T^{\square}$$

acting on $M^{\square}$. By how we chose $M^{\square}$ as a space of modular forms on a quaternion algebra $\varphi^{\square}$ will factor through the deformation ring with local conditions to give a map:

$$\varphi' : R^{\square} \to T^{\square}.$$

We will show this map is an isomorphism after inverting p. *Technical Aside*: To make everything work on the automorphic side one has to allow ramification at an auxiliary prime, this may cause the map φ' not to be surjective and so this has to be dealt with, but we won't worry about it here.

Observe that to show φ' is injective, it suffices to show that $R^{\square}$ acts faithfully on $M^{\square}$. Hence we forget T and focus on $M^{\square}$. We prove the following theorem:

Theorem 3.1. *The module $M^{\square}[1/p]$ is a finite projective (hence faithful) module over $R^{\square}[1/p]$. Further, $R^{\square}$ is finite over $\mathcal{O}[[y_1, \ldots, y_j]]$.*

Corollary 1. *The map*

$$\varphi'[1/p] : R^{\square}[1/p] \to T^{\square}[1/p]$$

is an isomorphism.

Note that by (H1) and (Q3) and some compatibilities, we can recover the action of $R^{\square}$ on $M^{\square}$ from any $(R^{\square}_{Q_n}, M^{\square}_{Q_n})$. This is the strategy we employ.

For each integer $n \geq 1$, with all the choices we have made we get a diagram:

$$\begin{array}{ccc} \mathcal{O}[[y_1, \ldots, y_j, T_1, \ldots, T_h]] & & \\ & \searrow & \\ B[[x_1, \ldots, x_g]] & \twoheadrightarrow & R^{\square}_{Q_n} \qquad \circlearrowleft M^{\square}_{Q_n}. \end{array}$$

There is absolutely no a priori relationship between the diagrams for different n. However, for the sake of exposition, assume there existed compatible maps between the diagrams. Once we explain the patching technique the same argument will go through.

Set $R_\infty = \lim_{\leftarrow} R^{\square}_{Q_n}$ and $M_\infty = \lim_{\leftarrow} R^{\square}_{Q_n}$. We get the following diagram:

$$\begin{array}{ccc} A = \mathcal{O}[[y_1, \ldots, y_j, T_1, \ldots, T_h]] & & \\ \downarrow & \searrow & \\ D = B[[x_1, \ldots, x_g]] & \longrightarrow & R_\infty \qquad \circlearrowleft M_\infty. \end{array}$$

where we pick a lift of the A-algebra structure on R_∞ to D, which we can do since its a power series ring.

Note that M_∞ is finite free over A. At each finite level, $M^{\square}_{Q_n}$ was finite free over $\mathcal{O}[[y_1, \ldots, y_j]][\Delta_{Q_n}]$, but in the limit, because we demanded higher and higher congruences for Δ_{Q_n}, we get freeness over the power series ring A.

Proposition 3.2. *Let $\psi : A \to D$ be a map of domains of the same dimension and let V be a D-module which is finite free as an A-module. Then, ψ is finite and if A and D are regular then V is a projective D-module.*

Proof of Theorem 3.1. We assume the proposition and deduce the theorem as a corollary (modulo actually constructing compatible maps). Both A and D from the diagram above are clearly domains. Proposition 2.4 tells us they have the same dimension.

Setting $V = M_\infty$ first, we get that the map $A \to B[[x_1, \ldots, x_g]]$ is finite. Since $B[[x_1, \ldots, x_g]]$ surjects onto R_∞, we get that R_∞ is finite over A. This remains true after quotienting both sides by the augmentation ideal $\mathfrak{a}$ to get back down to $R_\square$.

A is already regular but D may not be. However, (B3) says that

$$D[1/p] = B[[x_1, \ldots, x_g]][1/p] = (B \hat{\otimes}_{\mathcal{O}} \mathcal{O}[[x_1, \ldots, x_g]])[1/p]$$

is regular. The second part of the proposition applies to $A[1/p], D[1/p], M_\infty[1/p]$ so $M_\infty[1/p]$ is projective over $D[1/p]$ hence faithful. Since $D[1/p]$ acts through $R^\infty[1/p]$ the map between them must be injective, hence an isomorphism.

Thus, $M_\infty[1/p]$ is projective over $R^\infty[1/p]$ and this property descends through the quotient by $\mathfrak{a}$ to $(R^{\square}[1/p], M^{\square}[1/p])$. □

Proof of Prop 3.2. Each $d \in D$ acts on V and that action commutes with the action of A. This gives a map

$$D \to \mathrm{End}_A(V).$$

Let D' be the image of this map.

$$\begin{array}{ccc} D & \longrightarrow & \mathrm{End}_A(V) \\ \uparrow & \nearrow & \uparrow \\ A & \xrightarrow[finite]{} & D' \end{array}$$

It's clear that D' is finite over A so $\dim D' = \dim A = \dim D$. But D is a domain so any proper quotient of D has strictly smaller dimension so $D' = D$, and hence D is finite over A.

Now, we recall the Auslander-Buchsbaum Theorem (CRT Th'm 19.1): Let R be local Noetherian ring and M be finite module over R. Assume M has finite projective dimension. Then,

$$\mathrm{pd}_R(M) + \mathrm{depth}(M) = \mathrm{depth}(R).$$

Assume A and D regular so that V has finite projective dimension over D. To show V is projective, it suffices to show the inequality

$$\mathrm{depth}(V) \geq \mathrm{depth}(D) = \dim(D).$$

Take any regular sequence $a_1, \ldots, a_{\dim(A)}$ for V over A. The images of these in D form a regular sequence for V over D so we are done. □

4 Patching Datum

The key idea in patching datum is that the deformation rings and auxiliary rings are determined by their finite artinian quotients. This leads to a pigeonhole argument to find compatible maps between diagrams.

We unfortunately begin with more notation:

- $m_A^{(n)}$ is the ideal generated by nth powers for any complete local ring A
- $M_n := M_{Q_n}^{\square}, M_0 := M^{\square}, R_n := R_{Q_n}^{\square}, R_0 := R^{\square}$
- $s =$ rank of M_n over $\mathcal{O}[[x_1, \ldots, x_j]][\Delta_{Q_n}] = \mathcal{O}[[x_1, \ldots, x_j, T_1, \ldots, T_h]]/\mathfrak{b}_n$
- $r_m := smp^m(h+j)$
- $c_m := (\pi_E^m, x_1^{p^m}, \ldots, x_j^{p^m}, (T_1+1)^{p^m} - 1, \ldots, (T_h+1)^{p^m} - 1)$

Remark 4.1. For $m \leq n$, we have an inclusion of ideals $\mathfrak{b}_n \subset c_m$. This is simply because for $k \geq m$, $(T_i+1)^{p^k} - 1$ is divisible by $(T_i+1)^{p^m} - 1$.

Definition 4.2. A *patching datum* (D, L) of level m consists of:

1. A complete local noetherian ring D which is a B-algebra and such that $m_D^{(r_m)} = 0$ together with a D-module L which is finite free over $\mathcal{O}[[x_1, \ldots, x_j, T_1, \ldots, T_h]]/c_m$ of rank s;

2. A sequence of maps of complete local $\mathcal{O}$-algebras

$$\mathcal{O}[[x_1, \ldots, x_j, T_1, \ldots, T_h]]/c_m \to D \to R_0/(c_m R_0 + m_{R_0}^{(r_m)})$$

where the second map is a map of B-algebras;

3. A surjection $B[[y_1, \ldots, y_g]] \twoheadrightarrow D$;

4. And a surjection of $B[[y_1, \ldots, y_g]]$-modules

$$L \twoheadrightarrow M_0/c_m M_0.$$

This definition may seem arbitrary at first. However, the following two key properties illustrate the relevance of the definition.

Proposition 4.3. *For any $n \geq m$, we can construct a patching datum $(D_{m,n}, L_{m,n})$ of level m out of (R_n, M_n) by taking*

$$D_{m,n} = R_n/(c_m R_n + m_{R_n}^{(r_m)})L_{m,n} = M_n/c_m M_n.$$

For each fixed level m, this yields an infinite sequence of patching datum, one for each $n \geq m$.

Proof. First, $D_{m,n}$ is quotient of R_n and so is a complete local Noetherian ring which inherits a B-algebra structure from R_n as well as a surjection

$$B[[y_1, \ldots, y_g]] \twoheadrightarrow D_{m,n}$$

which we fixed earlier for R_n.

The desired sequence of maps comes from reducing

$$\mathcal{O}[[x_1, \ldots, x_j, T_1, \ldots, T_h]] \to R_n \to R_0$$

modulo c_m, $(c_m R_n + m_{R_n}^{(r_m)})$, and $(c_m R_0 + m_{R_0}^{(r_m)})$ respectively.

Input (H1) tells us the M_n is finite free over $\mathcal{O}[[x_1, \ldots, x_j, T_1, \ldots, T_h]]/\mathfrak{b}_n$. Since $\mathfrak{b}_n \subset c_m$, $L_{m,n} = M_n/c_m M_n$ is finite free over $\mathcal{O}[[x_1, \ldots, x_j, T_1, \ldots, T_h]]/c_m$ of the same rank. The surjective map

$$L_{m,n} \twoheadrightarrow M_0/c_m M_0$$

comes from reducing the map $M_n \twoheadrightarrow M_0$ modulo c_m.

It turns out the one non-trivial check is that $L_{m,n}$ is actually a module over $D_{m,n}$. Since M_n is an R_n-module $L_{m,n} = M_n/c_m M_n$ is an $R_n/c_m R_n$ - module. It suffices to show that $m_{R_n}^{(r_m)}$ acts trivially on $M_n/c_m M_n$.

Let $a \in m_{R_n}$ then a acts on $M_n = M_{Q_n}^{\square}$ via the Hecke algebra $T_{Q_n}^{\square}$. Consider the action of a on the quotient

$$M_n/(\pi_E, x_1, \ldots, x_j, T_1, \ldots, T_h)M_n = M_0/(\pi_E, x_1, \ldots, x_j)M_0$$

which is a finite $\mathbb{F}$ vector space of rank s. Since a lies in the maximal ideal of the Hecke algebra it acts as a nilpotent endomorphism hence

$$a^s M_n \subset (\pi_E, x_1, \ldots, x_j, T_1, \ldots, T_h)M_n.$$

A standard pigeonhole argument implies that

$$a^{sp^m(h+j)}M_n \subset (\pi_E, x_1^{p^m}, \ldots, x_j^{p^m}, T_1^{p^m}, \ldots, T_h^{p^m})M_n.$$

Raising to m, to get necessary power of π_E, in there, we conclude that

$$a^{sp^m(h+j)m}M_n \subset c_m M_n.$$

If you are struggling like me to keep track of all the exponents, the important point is that there is a fixed power of a which only depends on m which lands you in $c_m M_n$. This is not hard to see once you have $a^s M_n \subset (\pi_E, x_1, \ldots, x_j, T_1, \ldots, T_h)M_n$. □

Proposition 4.4. *There exist finitely many isomorphism classes of patching datum of level m.*

Proof. The number of elements in D is bounded above by the size of

$$B[[y_1, \ldots, y_g]]/m^{(r_m)}_{B[[y_1,\ldots,y_g]]}.$$

Also, L is free over finite ring. Its not hard to see from here that there are finitely many ways of putting the various structures on (D, L). □

Finally, we come to the salvage for our earlier passing to the limit argument. Consider the following arrangement of the data:

$(D_{1,1}, L_{1,1})$

$(D_{1,2}, L_{1,2})$ $(D_{2,2}, L_{2,2})$

$(D_{1,3}, L_{1,3})$ $(D_{2,3}, L_{2,3})$ $(D_{3,3}, L_{3,3})$

$(D_{1,4}, L_{1,4})$ $(D_{2,4}, L_{2,4})$ $(D_{3,4}, L_{3,4})$ $(D_{4,4}, L_{4,4})$

$(D_{1,5}, L_{1,5})$ $(D_{2,5}, L_{2,5})$ $(D_{3,5}, L_{3,5})$ $(D_{4,5}, L_{4,5})$ $(D_{5,5}, L_{5,5})$

$(D_{1,6}, L_{1,6})$ $(D_{2,6}, L_{2,6})$ $(D_{3,6}, L_{3,6})$ $(D_{4,6}, L_{4,6})$ $(D_{5,6}, L_{5,6})$ $(D_{6,6}, L_{6,6})$

The columns correspond to patching datum of increasing levels. In the first column, we can choose a infinite subsequence of isomorphic patching datum of level 1. Call it (D_1, L_1). In the second column consider the subsequence already chosen and pick a sub-subsequence all of whose entries at level 2 are isomorphic. Call it (D_2, L_2). Repeating this process, we get a sequence (D_i, L_i) of patching datum of level i such that the reduction to a lower level $(\tilde{D}_i, \tilde{L}_i) \cong (D_{i-1}, L_{i-1})$.

Taking the inverse limit, we get a pair (D_∞, L_∞) which one checks has the same properties as the R_∞, M_∞ considered in the previous section. By this remarkable process, we manage to piece together seemingly disconnected pieces of information to build a tower which with some clever commutative algebra proves our modularity lifting theorem.

5 Appendix A: Algebra Lemmas

Proposition 5.1. *Let $R \to R'$ be a local homomorphism of complete local Noetherian rings. Then, R' is flat over R if and only if $R'/m^n R'$ is flat over $R/m^n R$ for $n \geq 1$.*

Proof. The slight difficulty here is that we are not assuming R' is finite type over R. The forward implication is clear. We can check flatness on finite type modules so assume

$$0 \to K \hookrightarrow M$$

is an injective map of finite type R-modules. To check that

$$K \otimes_R R' \to M \otimes_R R'$$

is injective, we would like to use that

$$K \otimes_R R' \otimes_R R/m^n R \to M \otimes_R R' \otimes_R R/m^n$$

is exact because its isomorphic to

$$K/m^n K \otimes_{R/m^n R} R'/m^n R' \to M/m^n M \otimes_{R/m^n R} R'/m^n R'.$$

The injectivity of the unquotiented map is thus equivalent to the statement that

$$\cap m^n (K \otimes_R R') = 0$$

i.e. that $V = K \otimes_R R'$ is separated as an R-module. Since K finite-type, V is finite-type as a R' module so standard Nakayama says that for any ideal $I \subset R'$ contained in the maximal ideal $\cap I^n V = 0$. Since the homomorphism is local $mR' \subset m_{R'}$. □

Lemma 5.2. *Let R_1 and R_2 be complete local Noetherian $\mathcal{O}$-algebras. Let E be the fraction field of $\mathcal{O}$ and A be any local Artinian E-algebra. Then,*

$$\Psi : \mathrm{Hom}_E(R_1 \hat{\otimes}_{\mathcal{O}} R_2[1/p], A) \to \mathrm{Hom}_E(R_1[1/p], A) \times \mathrm{Hom}_E(R_2[1/p], A)$$

is a bijection.

Proof. Since p is invertible in A, we get

$$\mathrm{Hom}_E(R_1 \hat{\otimes}_{\mathcal{O}} R_2[1/p], A) = \mathrm{Hom}_{\mathcal{O}}(R_1 \hat{\otimes}_{\mathcal{O}} R_2, A).$$

Before completing, we have

$$\mathrm{Hom}_{\mathcal{O}}(R_1 \otimes_{\mathcal{O}} R_2, A) = \mathrm{Hom}_{\mathcal{O}}(R_1, A) \times \mathrm{Hom}_{\mathcal{O}}(R_2, A)$$

so the only question is does a homomorphism $f : R_1 \otimes_{\mathcal{O}} R_2 \to A$ extend to the completion. If it does, it does so uniquely.

Write $R_1 \cong \mathcal{O}[[x_1, \ldots, x_n]]/(g_1, \ldots, g_r)$ and $R_2 \cong \mathcal{O}[[y_1, \ldots, y_m]]/(h_1, \ldots, h_s)$. Let f_1, f_2 be the induced maps $R_1 \to A, R_2 \to A$ respectively. Now, A has both a reduction map $A \to E$ and section $E \to A$. We know from Brian's Lecture 6 that under the reduction maps $f_1(x_i)$ and $f_2(y_j)$ map to elements d_i, e_j respectively in the maximal ideal of $\mathcal{O}$. Considering d_i and e_j as elements of A under the section map, we see that $f_1(x_i - d_i) \in m_A$ and similarly $f_2(y_j - e_j) \in m_A$.

Now, let k be an integer such that $m_A^k = 0$. Then, its clear that

$$(x_1 - d_1, \ldots, x_r - d_r)^k \subset \ker f_1 \text{ and } (y_1 - e_1, \ldots, y_s - e_s)^k \subset \ker f_2.$$

Hence, the morphism f factors through $R_1/(x_1 - d_1, \ldots, x_r - d_r)^k \otimes_{\mathcal{O}} R_2/(y_1 - e_1, \ldots, y_s - e_s)^k$. Both quotients, however, are now polynomial rings over $\mathcal{O}$ so the completed tensor product is the same as the ordinary tensor product and so f trivially extends to the completion. $\square$