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 representation 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 GL(2, Q_p) (or more generally, GL(2, F) with F/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 about the Langlands correspondence (both local and global) that we may not have covered.

1.1. Admissible representations. Let F/Q_p be a finite extension and let G be the group GL(2, F). Fix an algebraically closed field K of characteristic zero (one always takes K to be the complex numbers or the closure of some Q_p). 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 \( \alpha \) and \( \beta \) 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

\[
    f \left( \begin{pmatrix} a & x \\ b & y \end{pmatrix} g \right) = \alpha(a)\beta(b) \left| \frac{a}{b} \right|^{1/2} f(g)
\]

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 \( \cdot \cdot \) or \( \cdot \cdot^{-1} \) then V is irreducible. Here \( \cdot \cdot \) is the norm character of \( F^\times \), which takes \( a \in F^\times \) to \( q^{-\text{val}_a} \) 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 \( \cdot \cdot \) or \( \cdot \cdot^{-1} \) 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. Precisely, say \( \alpha\beta^{-1} = \cdot \cdot \) and write \( \alpha = \gamma \cdot \cdot^{1/2} \) and \( \beta = \gamma^{-1} \cdot \cdot^{-1/2} \). Then \( V(\alpha, \beta) \) has a unique irreducible subrepresentation \( \text{St}(\gamma) \) which is infinite dimensional. The quotient \( V(\alpha, \beta)/\text{St}(\gamma) \) is one dimensional and \( G \) acts on it through the character \( g \mapsto \gamma(\det g) \). Write \( \text{St} \) in place of \( \text{St}(\gamma) \) where \( \gamma \) is the trivial character. The representation \( \text{St} \) is called the Steinberg representation. One has \( \text{St}(\gamma) = \text{St} \otimes \gamma \).

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 representations, 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 $\text{St} \otimes \gamma$ of the Steinberg representation $\text{St}$.
• 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 $\text{GL}(2, \mathcal{O}_F)$. It is a theorem that $V$ is unramified if and only if it is a principal series of the form $V(\alpha, \beta)$ with $\alpha$ and $\beta$ 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 $\gamma$ 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

$$0 \to I_F \to \text{Gal}(\overline{F}/F) \cong \mathbb{Z} \to 0$$

where $I_F$ is the inertia subgroup of the Galois group. The Weil group of $F$ is by definition the subgroup of $\text{Gal}(\overline{F}/F)$ given by $\text{val}^{-1}(\mathbb{Z})$. 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

$$gNg^{-1} = q^\text{val} g N$$

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 $g \mapsto q^\text{val} g$.

The collection of all Weil-Deligne representations forms a category and this category 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:

$$\left\{ \begin{array}{c}
\text{Weil-Deligne representations} \\
\text{with coefficients in } \mathbb{Q}_\ell
\end{array} \right\} \leftrightarrow \left\{ \begin{array}{c}
\text{Continuous Frobenius semi-simple representations of } W_F \text{ on } \mathbb{Q}_\ell \text{ vector spaces}
\end{array} \right\}$$

**Sketch of proof.** Let $(V, N)$ be a Weil-Deligne representation. Let $\rho$ denote the action of $W_F$ on $V$. Define a new representation $\rho'$ of $W_F$ on $V$ by

$$\rho'(\text{Frob}^n g) = \rho(\text{Frob}^n g) \exp(Nt_\ell(g)).$$

Here Frob $\in W_F$ is a fixed Frobenius element, $g$ is an element of the inertia subgroup $I_F$ of $W_F$ and $t_\ell : I_F \to \mathbb{Z}_\ell$ is the tame $\ell$-adic character. One easily verifies that $\rho'$ 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-Deligne 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 $\gamma$ of $W_F$); $N$ kills $W(1)$ and maps $W$ isomorphically onto $W(1)$. 

\[ \square \]
We can now state a version of the local Langlands correspondence for $GL(2)$.

**Theorem 1.4.** There is a natural bijection

\[
\begin{align*}
\{ \text{Irreducible admissible representations of } GL(2,F) \text{ over } K \} & \leftrightarrow \{ \text{Two dimensional Weil-Deligne representations with coefficients in } K \}.
\end{align*}
\]

Under this 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 $\alpha'$ and $\beta'$ correspond to $\alpha$ and $\beta$ 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 $GL(2,F)$. 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 $\ell$-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 $A_F$ be its adele ring. An **automorphic form** on $GL(2)$ over $F$ is a function $f : GL(2, A_F) \to \mathbb{C}$ satisfying a number of properties, the most important of which is that it is invariant on the left under $GL(2,F)$. The set of all automorphic forms forms a vector space $\mathcal{A}_F$. This vector space carries an action of $GL(2, A_F')$ by right translation. Furthermore, the Lie algebra and the maximal compact of $GL(2, F_\infty)$ act on $\mathcal{A}_F$ (that is, $\mathcal{A}_F$ is a Harish-Chandra module for $GL(2, F_\infty)$). (The full group $GL(2, F_\infty)$ does not act on $\mathcal{A}_F$ as it destroys the $K$-finiteness condition.)

An automorphic representation of $GL(2, A_F)$ is something of the form $\pi_f \otimes \pi_\infty$ where $\pi_f$ is an irreducible admissible representation of $GL(2, F_f)$ and $\pi_\infty$ is an irreducible Harish-Chandra module of $GL(2, F_\infty)$ such that $\pi_f \otimes \pi_\infty$ is equivalent to a submodule of $A_F$. There is a certain condition called cuspidal that one can impose on automorphic forms. The set of all cuspidal forms forms a vector subspace $\mathcal{A}_F^c$ of $\mathcal{A}_F$ which is stable under the various actions of pieces of $GL(2, A_F)$. An automorphic representation is cuspidal if it appears inside this cuspidal space.

Say for the moment that $F = \mathbb{Q}$. As we have discussed earlier in the semester, classical modular eigenforms correspond bijectively to automorphic representations $\pi$ for which $\pi_\infty$ is a discrete series representation. More precisely, say $f$ is a newform of level $N$ and weight $k$ and let $\pi$ be the corresponding automorphic representation. We can then write $\pi = \pi_f \otimes \pi_\infty$ and further decompose $\pi_f$ as a restricted tensor product $\otimes \pi_p$, where $\pi_p$ is an irreducible admissible representation of $GL(2, \mathbb{Q}_p)$. The Harish-Chandra module $\pi_\infty$ is completely determined by the weight $k$. For primes $p$ not dividing the level, $\pi_p$ is an unramified representation of $GL(2, \mathbb{Q}_p)$. As we have seen, such representations are determined by two numbers; the representation $\pi_p$ 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 $\alpha$ and $\beta$ of $\mathbb{Q}_p^\times$ such that $\pi_p$ is equivalent to $V(\alpha, \beta)$.)

For primes $p$ dividing $N$ the representation $\pi_p$ is not unramified. I imagine that it is possible to determine $\pi_p$ 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 $F \neq \mathbb{Q}$ 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 Hilbert 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 $b^n/\Gamma_i$ where $b$ is the upper half plane and the $\Gamma_i$ are certain arithmetic groups. The proper analogue of a modular form is then a tuple $(f_i)$ where $f_i$ is a function on $b^n$ invariant under $\Gamma_i$. 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 switching to the representation theoretic perspective.

**Key points:** (1) Classical modular forms correspond to automorphic representations of $GL(2, A_\mathbb{Q})$ 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

$$\rho_{f,\ell} : G_{\mathbb{Q}} \to \text{GL}_2(\overline{\mathbb{Q}}_{\ell})$$

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 $\rho_{f,\ell}(\text{Frob}_p)$ 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 $\pi$ be an automorphic representation of $\text{GL}(2, \mathbb{A}_F)$ such that $\pi_{\infty}$ is a discrete series representation. Then there is a Galois representation

$$\rho_{\pi,\ell} : G_{F} \to \text{GL}_2(\overline{\mathbb{Q}}_{\ell})$$

which satisfies and is uniquely determined by the following property: if $v$ is a place of $F$ which does not lie above $\ell$ then $\rho_{\pi,\ell}|G_{F,v}$ corresponds to $\pi_v$ 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 $\pi_{\infty}$ 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 $\mathbb{Q}$. However, even for $F = \mathbb{Q}$ it is a stronger result: it specified the local Galois representation everywhere except at $\ell$ in terms of the corresponding local component of the automorphic representation. The local Galois representation at $\ell$ 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 $\ell$ (coming from $\ell$-adic Hodge theory). This has basically been proved for $F = \mathbb{Q}$ but is still open for all other $F$. The most critical intermediate result in the proof for $F = \mathbb{Q}$ is a modular lifting theorem; we will prove such a theorem in this seminar.

Key point: Given an automorphic representation $\pi$ of a totally real number field which is discrete series at infinity, there is a corresponding Galois representation $\rho_{\pi,\ell}$. (Or rather, one for each $\ell$.) The restriction of $\rho_{\pi,\ell}$ to a decomposition group away from $\ell$ corresponds to the local component of $\pi$ 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 $\text{GL}(2)$ 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:

$$\{\text{Automorphic representations of } G' \text{ which are essentially square integrable at all places where } D \text{ ramifies} \} \leftrightarrow \{ \text{Automorphic representations of } G \}$$

(An irreducible admissible representation of $\text{GL}(2, F_v)$ is essentially square integrable if it is a twist of the Steinberg or supercuspidal, i.e., not principal series.) Furthermore, if $\pi'$ is an automorphic representation of $G'$ and $\pi$ the corresponding automorphic representation of $G$ then $\pi_v$ is determined completely by $\pi'_v$. Two special cases: (1) if $D$ splits at $v$ and we identify $D_v$ with $M_2(F_v)$ then $\pi'_v$ is identified with $\pi_v$; (2) if $\pi'_v$ is the trivial representation then $\pi_v$ 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 $(D \otimes \mathbb{A}_F) \times$ let $S_2(U)$ denote the space of all functions

$$D^\times \backslash (D \otimes \mathbb{A}_F) \times / U \to \mathbb{C}.$$ 

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 Eisenstein series 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 $\pi$ be an automorphic representation of $GL(2, \mathbb{A}_F)$ with $F$ a number field, such that $\pi_\infty$ 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 $\pi'$ of $GL(2, \mathbb{A}_{F'})$. The automorphic representation $\pi'$ 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

$$BC : \left\{ \text{Automorphic representations of } GL(2, \mathbb{A}_F) \right\} \to \left\{ \text{Automorphic representations of } GL(2, \mathbb{A}_{F'}) \right\}$$

such that if $\pi' = BC(\pi)$ then: (1) the local component $\pi'_v$ can be computed in terms of $\pi_v$; (2) if $\pi_\infty$ is discrete series then so is $\pi'$ 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 $GL(2, F_v)$ to those of $GL(2, F'_v)$. In fact, the meaning of (1) in the above theorem is precisely that $\pi'_v = BC(\pi_v)$. 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 $\alpha' \boxplus \beta'$ where $\alpha'$ and $\beta'$ correspond to $\alpha$ and $\beta$ under class field theory. Restricting this to $G_{F_v}$ we simply get $\alpha'|_{G_{F_v}} \boxplus \beta'|_{G_{F_v}}$. Going the other way under local Langlands, this corresponds to the principal series $V(\alpha'', \beta'')$ where $\alpha''$ and $\beta''$ 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

$$BC(V(\alpha, \beta)) = V(N^* \alpha, N^* \beta).$$

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 $\pi$ is any irreducible admissible representation of $GL(2, F_v)$ then one can find an extension $F'_v/F_v$ such that $BC(\pi)$ is either unramified or Steinberg. Any base change of Steinberg is still Steinberg, however.

The above local discussion has the following global application (when combined with some global class field theory). Given an automorphic representation $\pi$ of $GL(2, \mathbb{A}_F)$ there exists a finite solvable Galois
extension $F'/F$ such that the base change of $\pi$ 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 $\rho : G_F \to \text{GL}_2(\mathbb{Q}_p)$ be a Galois representation. Assume that there exists a finite, totally real, solvable extension $F'/F$ and a parallel weight 2 automorphic representation $\pi'$ of $\text{GL}(2, \mathbb{A}_F')$ such that $\rho|_{G_{F'}} = \rho_{\pi',p}$ and both are irreducible. Then there exists a parallel weight 2 automorphic representation $\pi$ of $\text{GL}(2, \mathbb{A}_F)$ such that $\rho = \rho_{\pi,p}$.

In other words: if $\rho$ becomes modular over a solvable extension then $\rho$ 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 checking over a unramified or Steinberg at all places. One cannot get rid of Steinbergs through base change, however.

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 $F'/\mathbb{Q}_p$ be a finite extension. We say that a Galois representation $\rho : G_F \to \text{GL}_2(\mathbb{Q}_p)$ is ordinary if it is of the form

\[
\begin{pmatrix}
\alpha \chi_p & * \\
0 & \beta
\end{pmatrix}
\]

where $\alpha$ and $\beta$ are finitely ramified characters, and, as always, $\chi_p$ denotes the $p$-adic cyclotomic character. (One could allow for more general definitions of ordinary, replacing $\chi_p$ by $\chi_p^n$: for now we will stick with this one.) Let $E/F$ be an extension over which $\alpha$ and $\beta$ become unramified. The representation $\rho|_{I_E}$ is an extension of the trivial representation by $\chi_p$ and so defines an element of $H^1(I_E, \mathbb{Q}_p(\chi_p))$, which is identified with $\mathbb{Q}_p \otimes (E^{un})^\times$ 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 $\rho$ is potentially crystalline if this class belongs to $\mathbb{Q}_p \otimes \Omega_{E^{un}}$. This is independent of the choice of $E$.

Now let $F/\mathbb{Q}$ be a finite totally real extension. Recall that a representation $\rho : G_F \to \text{GL}_2(\mathbb{Q}_p)$ 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 $\rho : G_F \to \text{GL}_2(\mathbb{Q}_p)$ be an odd, finitely ramified representation such that $\overline{\rho}|_{G_{F}(\mathbb{Q}_p)}$ is absolutely irreducible and $\rho$ is potentially crystalline and ordinary at all places above $p$. Assume that there exists an automorphic representation $\pi$ of $\text{GL}(2, \mathbb{A}_F)$ such that $\rho_{\pi,p}$ is potentially crystalline and ordinary at all places above $p$ and $\overline{\pi}_p = \overline{\rho}$. Then there exists an automorphic representation $\pi'$ 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 $\rho$ is of the form

\[
\begin{pmatrix}
\alpha \chi_p & * \\
0 & \beta
\end{pmatrix}
\]

at each place above $p$. By making a solvable base change, we can reduce to the case where $\alpha$ and $\beta$ are unramified. Even more drastically, we can make a solvable base change to reduce to the case where $\overline{\rho}|_{G_F}$ 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 $F/\mathbb{Q}$ 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 representation
such that $\mathfrak{p} = \mathfrak{p}_{\pi,p}$. In the proof, however, we need it to be the case that $\rho$ and $\rho_{\pi,p}$ are potentially unramified at the same set of places. This need not be the case for the $\pi$ we have. Of course, we are free to replace $\pi$ with another form $\pi'$ such that $\mathfrak{p}_{\pi,p} = \mathfrak{p}_{\pi',p}$, that is, one that is congruent to $\pi$ modulo $p$ (while still maintaining the other hypotheses). So the question is: given $\pi$ as in the theorem, can we find a congruent $\pi'$ such that $\rho_{\pi',p}$ and $\rho$ 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 $\pi$ 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 certain universal deformation ring of the Galois representation $\rho$ 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.