AUTOMORPHIC REPRESENTATIONS OF INNER FORMS OF $GL(2)$

M. Harris, Luminy 2007

1. Local theory of representations of GL(2)

The irreducible admissible representations of $GL(2, F)$ have been divided into three classes, for convenience, since the time of Jacquet-Langlands (and even before). These are the *principal series*, the *supercuspidal representations*, and the *Steinberg* (or special) representations. This classification is incomplete; it does not include the finite dimensional representations, which are (almost) irrelevant to the global theory, as we'll see below. I will state a lemma, which I leave as an exercise:

Lemma 1.1.1. *Let π be a finite-dimensional smooth irreducible complex representation of $GL(2, F)$. Then $\dim \pi = 1$ and there is a continuous (locally constant) character $\chi : F^\times \to \mathbb{C}^\times$ such that $\pi = \chi \circ \det$.*

Remark. In the above lemma, and in most of what I will be writing in this section, it is immaterial that the coefficient field be $\mathbb{C}$: it works just as well for smooth representations with coefficients in any algebraically closed field of characteristic zero.

Now for the three main classes:

1.1.2 Principal series. Let (χ_1, χ_2) be an ordered pair of characters of $F^\times$. Let $G = GL(2, F)$, $B \subset G$ the upper-triangular Borel subgroup, $B = A \cdot N$, where

$$A = \{\begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}\},\ N = \{\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}\}.$$

Define

$$I(\chi_1, \chi_2) = \{f : G \to \mathbb{C} \mid f(ang) = \chi_1(a_1)|a_1|^{\frac{1}{2}}\chi_2(a_2)|a_2|^{-\frac{1}{2}} \cdot f(g)\}.$$

This is a (normalized) induced representation and G acts on $I(\chi_1, \chi_2)$ by right translation:

$$r(g)f(h) = f(hg).$$

Those half-powers of the norm guarantee that

$$I(\chi_1, \chi_2) \xrightarrow{\sim} I(\chi_2, \chi_1)$$

almost always. More precisely

Typeset by $\mathcal{A}_\mathcal{M}\mathcal{S}$-TEX

Proposition 1.1.2.1. *(a) $I(\chi_1,\chi_2)$ is irreducible unless $\chi_1/\chi_2 = |\bullet|^{\pm 1}$. (b) If $\chi_1/\chi_2 \neq |\bullet|^{\pm 1}$ then $I(\chi_1,\chi_2) \xrightarrow{\sim} I(\chi_2,\chi_1)$ as irreducible admissible representations of $GL(2,F)$. (c) If $\chi_1/\chi_2 \neq |\bullet|^{\pm 1}$ then $I(\chi_1,\chi_2)^\vee \xrightarrow{\sim} I(\chi_1^{-1},\chi_2^{-1})$.*

Let $K = GL(2,\mathcal{O}) \subset G$, where $\mathcal{O}$ is the integer ring of F. Suppose χ_1 and χ_2 are *unramified*. Then $I(\chi_1,\chi_2)$ contains a canonical K-invariant vector f_0 defined by

$$f_0(k) = 1 \forall k \in K.$$

This is unique because of the Iwasawa decomposition $G = B \cdot K$. For any ring A one can define a local Hecke algebra $H_A(G,K)$ as the convolution algebra of compactly supported functions on G that are right-and left-invariant under K-translation. The algebra $H_A(G,K)$ is commutative and if A is a $\mathbb{Q}$-algebra there are canonical generators

$$T = K \cdot \begin{pmatrix} \varpi & 0 \\ 0 & 1 \end{pmatrix} \cdot K, \; RK \cdot \begin{pmatrix} \varpi & 0 \\ 0 & \varpi \end{pmatrix} \cdot K,$$

where ϖ is a uniformizer in $\mathcal{O}$, so that

$$H_A(G,K) = A[T,R,R^{-1}].$$

The convolution algebra $H(G)$ of all locally constant compactly supported functions operates on any (admissible) representation of G, and the function f_0 is an eigenvector for the subalgebra $H(G,K)$, with

$$Tf_0 = q^{\frac{1}{2}}(\chi_1(\varpi)+\chi_2(\varpi))f_0, \; Rf_0 = \chi_1(\varpi)\chi_2(\varpi)f_0, \tag{1.1.2.2}$$

where q is the order of the residue field. Any irreducible representation with a K-fixed vector is determined up to isomorphism by its T and R-eigenvalues; such a representation is called *spherical.*

If $F = \mathbb{Q}_p$ then T is the classical Hecke operator $T(p)$, and $R = T(p,p)$. The classical theory of Hecke operators is completely subsumed by the theory of spherical principal series.

1.1.2 Steinberg representations.

Let $\chi : F^\times \to \mathbb{C}^\times$ be a smooth character. Let

$$\chi_1 = \chi \cdot |\bullet|^{-\frac{1}{2}}, \chi_2 = \chi \cdot |\bullet|^{\frac{1}{2}}.$$

In that case it is easy to see that the function $f_\chi(g) = \chi(det(g)) \in I(\chi_1,\chi_2)$. There is a short exact sequence of admissible representations

$$0 \to \mathbb{C}f_\chi \to I(\chi_1,\chi_2) \to St(\chi) \to 0$$

where G acts by $\chi \circ \det$ on $\mathbb{C}f_\chi$ and $St(\chi)$ is irreducible; these are the Steinberg representations. When χ is trivial, one calls $St(1)$ *the* Steinberg representation.

The representation $I(\chi_2,\chi_1)$ is again reducible and has $St(\chi)$ as subspace, $\chi \circ \det$ as quotient.

1.1.3 Supercuspidal representations.

Sections 1.1.1 and 1.1.2 complete the analysis of principal series. Any irreducible representation of G that has not yet been mentioned is called *supercuspidal.*

Theorem 1.1.3.1. *Let π be an irreducible admissible representation of $GL(2, K)$. The following are equivalent:*

(i) π is supercuspidal

(ii) For all χ_1, χ_2, $Hom(\pi, I(\chi_1, \chi_2)) = 0$.

(iii) The Jacquet modules of π vanish

Iiv) The matrix coefficients of π are compactly supported modulo translation by the center Z of G.

This doesn't help identify supercuspidal representations, and even now, long after they have been classified, there is no simple way to present the supercuspidals.

Definition 1.1.3.2. *An irreducible admissible representation π of G is called* discrete series *(or* square-integrable*) if it is either Steinberg or supercuspidal.*

The term "discrete series" comes from the theory of harmonic analysis on G; these are the representations that contribute to the Plancherel measure on G non-trivially (when they are unitary). Equivalently, they occur discretely in the Hilbert decomposition of $L_2(G/Z)$ under right and left-translation by G.

1.2. Parametrization of irreducible admissible representations of $GL(2, F)$.

Let $\mathcal{A} = \mathcal{A}_{2,F}$ denote the set of equivalence classes of irreducible admissible representations of $GL(2, F)$. There are several uniform parametrizations of $\mathcal{A}$.

1.2.1 Distribution characters. If π is admissible and $\phi \in H(G)$ then $\phi : \pi \to \pi$ has finite image, hence, one can define its $Tr(\phi; \pi)$. The linear functional

$$Tr_\pi : H(G) \to \mathbb{C}; Tr_\pi(\phi) = Tr(\phi; \pi)$$

is a distribution. Fix a Haar measure dg on G (usually the one with measure 1 on K).

Theorem 1.2.1.1. *The distribution Tr_π that is represented by a locally integrable function, denoted χ_π, on G:*

$$Tr_\pi(\phi) = \int_G \chi_\pi(g)\phi(g)dg.$$

The function χ_π is **invariant***:*

$$\chi_\pi(hgh^{-1}) = \chi_\pi(g) \forall h \in G.$$

Of course χ_π depends on the choice of dg. The function χ_π usually has singularities on G but is continuous on the subset G^{reg} of regular elements: $g \in G$ is regular if and only if its characteristic polynomial has distinct roots (in which case it is certainly semisimple.

Definition 1.2.1.2. *A regular semisimple $g \in G$ is* elliptic *if its characteristic polynomial is irreducible over F. The set of regular elliptic elements is denoted G^{re}.*

Proposition 1.2.1.3. *Let π be irreducible admissible and infinite dimensional. If $\chi_\pi(g) \neq 0$ for some $g \in G^{re}$, then π is discrete series, and in that case $\chi_\pi(g) = 0$ if $g \in G^{reg} \setminus G^{re}$.*

The distribution character χ_π determines π up to isomorphism (up to semisimplification if π is an admissible representation of finite length). Moreover, linear independence of characters is valid for irreducible admissible representations. There are explicit formulas for χ_π when π is principal series and these are useful in the trace formula, but we will not describe them.

1.2.2. Langlands parametrization.

The most conceptually useful parametrization of $\mathcal{A}$. Let $\mathcal{G} = \mathcal{G}_{2,F}$ denote the set of equivalence classes of (Frobenius-semisimple) two-dimensional representations of the Weil-Deligne group WD_F. Let

$$r_F : (WD_F)^{ab} \xrightarrow{\sim} F^\times$$

be the (normalized) reciprocity isomorphism.

1.2.2.1. Local Langlands correspondence (Kutzko). *There is a bijection*

$$\mathcal{L} : \mathcal{A}_{2,F} \to \mathcal{G}_{2,F}$$

with the following natural properties, among others;

(i) *If $I(\chi_1, \chi_2)$ is irreducible, then*

$$\mathcal{L}(I(\chi_1, \chi_2)) = \chi_1 \circ r_F \oplus \chi_2 \circ r_F.$$

(ii) *π is supercuspidal if and only if $\mathcal{L}(\pi)$ is irreducible.*
(iii) *$\mathcal{L}(St(\chi)) = \chi \otimes Sp(2)$ where $Sp(2)$ is the trivial extension of the non-trivial 2-dimensional representation of the unipotent $N \in WD_F$.*
(iv) *$\mathcal{L}$ preserves local L and ε factors.*
(v) *$\det \mathcal{L}(\pi)) = r_F \circ \xi_\pi$, where ξ_π is the central character of π.*

A word about (iv): In what follows, I take the attitude that everything concerning representations of WD_F is understood. In particular, one could define local L and ε factors of irreducible admissible representations of $GL(2, F)$ by (iv). Of course, these factors were defined by Jacquet-Langlands and coincide with Hecke's local Euler factors of elliptic modular forms in the classical setting, and (iv) is the non-trivial theorem that the two definitions are compatible.

Example (Iwahori-stable vectors). Let $I \subset K$ be the Iwahori subgroup, generalizing the classical $\Gamma_0(p)$:

$$I = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2, \mathcal{O}) \mid c \in \mathfrak{m}_\mathcal{O}\}$$

where $\mathfrak{m}_\mathcal{O} \subset \mathcal{O}$ is the maximal ideal. We are interested in classifying irreducible representations π of G with vectors $v \in \pi$ on which I acts by a character. Let $I(1)$ be the subgroup of I as above with $a, d \equiv 1 \pmod{()}\mathfrak{m}_\mathcal{O}$). The classification is simple.

Proposition 1.2.2.2. *Let π be an irreducible admissible representation of G which contains a vector on which I acts by a character trivial on $I(1)$. Then π is one of the following:*

(a) *π is an irreducible spherical principal series $I(\chi_1, \chi_2)$, with χ_1 and χ_2 unramified, and $\dim \pi^K = 1$, $\dim \pi^I = 2$. In this case $\mathcal{L}(\pi) = \chi_1 \circ r_F \oplus \chi_2 \circ r_F$.*

(b) *π is 1-dimensional of the form $\chi \circ \det$, with χ unramified. In this case $\mathcal{L}(\pi) = \chi \circ r_F \oplus \chi \cdot |\bullet|_F \circ r_F$.*

(c) *$\pi = St(\chi)$, with χ unramified . In this case $\dim \pi^I = 1$, and $\mathcal{L}(\pi)$ is as in 1.2.2.1 (iii).*

(d) *$\pi = St(\chi)$, with χ ramified of conductor $\mathfrak{m}_{\mathcal{O}}$. In this case $\dim \pi^{I(1)} = 1$ and $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in I$ acts on $\pi^{I(1)}$ by $\chi(ad)$; $\mathcal{L}(\pi)$ is as in 1.2.2.1 (iii).*

(d) *π is a principal series $I(\chi_1, \chi_2)$ with one or both of χ_1, χ_2 ramified of conductor $\mathfrak{m}_{CO}$, the other possibly unramified. In this case $\pi^{I(1)} = 2$ is the sum of two one-dimensional eigenspaces for $I/I(1)$ with eigenvalues*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in I \mapsto \chi_1(a)\chi_2(d),\ \ \chi_2(a)\chi_1(d)$$

and $\mathcal{L}(\pi)$ is as in 1.2.2.1 (i).

1.2.3. Local base change and automorphic induction.

Let F'/F be a finite extension. The inclusion $WD_{F'} \subset WD_F$ defines a natural restriction map

$$\mathcal{G}_{2,F} \to \mathcal{G}_{2,F'},$$

hence by the local Langlands parametrization a natural base change map

$$BC = BC_{F'/F} : \mathcal{A}_{2,F} \to \mathcal{A}_{2,F'}.$$

When F'/F is a cyclic Galois extension of prime degree, the map $BC_{F'/F}$ coincides with Langlands' construction and is determined by an explicit relation between χ_π and $\chi_{BC\pi}$. The generalization of Langlands' base change construction, due to Arthur and Clozel, is a crucial ingredient in every known construction of the local Langlands correspondence for general $GL(n)$, For $n = 2$ an alternative construction not explicitly referring to the distribution character is developed in the book of Bushnell-Henniart.

Let F'/F be quadratic, and define $\mathcal{G}_{1,F'}$, resp. $\mathcal{A}_{1,F}$, to be the set (group) of continuous characters of $WD_{F'}$, resp. $F'^{,\times}$. Thus local class field theory identifies $\mathcal{G}_{1,F'} \xrightarrow{\sim} \mathcal{A}_{1,F'}$. Induction defines a natural map

$$\mathcal{G}_{1,F'} \to \mathcal{G}_{2,F},$$

hence by the local Langlands parametrization a natural map called (local) **automorphic induction**:

$$AI_{F'/F} : \mathcal{A}_{1,F'} \to \mathcal{A}_{2,F}.$$

Proposition 1.2.3.1. *If F is of residue characteristic $p > 2$ then every supercuspidal representation of $GL(2, F)$ is in the image of $AI_{F'/F}$ for some quadratic extension F'/F.*

This is because every irreducible two-dimensional representation of the Weil group of F is tame when $p \neq 2$, and it is easy to see that it is therefore induced from a character of the Galois group of a quadratic extension.

However, there is an alternative explicit construction of the automorphic induction map, using the Weil representation; it is a local version of the theory of theta functions of binary quadratic forms. Proposition 1.2.3.1 was first proved by means of this construction. Kutzko's theorem therefore really concerns the case $p = 2$.

1.2.4. Other realizations.

I briefly mention two alternative approaches to $\mathcal{A}_{2,F}$. One is purely group-theoretic and is based on the structure of the lattice of open compact subgroups of G. This is called the theory of *types*, which parametrizes the supercuspidal representations in $\mathcal{A}_{2,F}$ in terms of explicit representations of K. For $GL(2, F)$ this was completely worked out by Kutzko, and the (much more difficult) classification for $GL(n, F)$ is due to Bushnell and Kutzko. The extension of this technique to other reductive p-adic groups is an active area of research, and there has been considerable progress in recent years for classical groups. The theory of types is the most useful parametrization for studying congruences among supercuspidal representations and between supercuspidal representations and others (cf. the modular local Langlands correspondence of Vignéras). It may be mentioned in some of the later talks.

The second construction of supercuspidal representations of $GL(2, F)$ is by means of arithmetic geometry – indeed, rigid geometry. This was first considered by Deligne and was completely developed by Carayol (for F of characteristic zero). All discrete series representations of $GL(2, F)$ can be realized on a space of vanishing cycles attached to the degeneration of one-dimensional height 2 formal groups (more generally formal $\mathcal{O}$-modules) with (Drinfeld) level structure. This same space of vanishing cycles is important in connection with Ribet's work on congruences of modular forms of different levels, and may be mentioned in later talks.

1.3. Classification of irreducible representations of $GL(2, \mathbb{R})$.

1.3.1. Harish-Chandra modules.

Now $G = GL(2, \mathbb{R})$, and $\mathfrak{g} = Lie(G) = M(2, \mathbb{R})$; $K = \mathbb{R}^{\times} \cdot O(2) \subset G$, $\mathfrak{k} = Lie(K) = Lie(\mathbb{R}^{\times} \cdot SO(2))$. An *admissible representation* of $GL(2, \mathbb{R})$, or Harish-Chandra module, is not a representation of $GL(2, \mathbb{R})$ at all, but rather is a (generally infinite-dimensional) module V for the universal enveloping algebra $U(\mathfrak{g}_{\mathbb{C}})$. One assumes the action of $\mathfrak{k}$ on V integrates to an action of $\mathbb{R}^{\times} \cdot SO(2)$ that extends to an action of K that decomposes V as a sum of *finite-dimensional* isotypic components for K. The actions of K and $\mathfrak{g}_{\mathbb{C}}$ are assumed to be compatible in a natural way. Harish-Chandra modules form an abelian category with excellent properties.

1.3.2. Holomorphic representations.

For purposes of arithmetic, we only need to consider the kinds of Harish-Chandra modules π that arise as archimedean local components of an automorphic representation Π of $GL(2, F)$, where F is now a totally real number field and Π is associated to a holomorphic Hilbert modular newform. Let $k > 0$ be an integer, which will be the weight of the modular form, and let $s \in \mathbb{C}$. There is a unique irreducible representation of $GL(2, \mathbb{R})$ (i.e., Harish-Chandra module), denoted $\pi_{k,s}$, that "contains a holomorphic vector of weight k" (i.e., arises as the local component of the adelic automorphic representation attached to a Hilbert modular form of weight k) and such that the scalar $t \in \mathbb{R}^\times$, $t > 0$, acts as t^s. In the standard normalization, we take $s = 2 - k$. This is compatible with the standard normalization at nonarchimedean places in the following sense: if π is a cuspidal automorphic representation of $GL(2)$ (see §2 below) with $\pi_\infty = \pi_{k,2-k}$, corresponding to a modular form of weight k, then the usual local Euler factor at an unramified prime p

$$1 - a_p p^{-s} + p^{k-1-2s})^{-1}$$

is identified with the Jacquet-Langlands local Euler factor $L(s - \frac{1}{2}, \pi_p)$ and $L(s - \frac{1}{2}, \pi_\infty)$ gives the usual Γ-factor for holomorphic modular forms of weight k. See also (3) of §3.

One sometimes uses the so-called unitary normalization, in which $s = 0$; this allows us to write the functional equation relating $L(s, \pi)$ and $L(1 - s, \bar{\pi})$.

There is also an automorphic induction map, where notation is as in 1.2.3:

$$\mathcal{A}_{1,\mathbb{C}} \to \mathcal{A}_{2,\mathbb{R}},$$

and $\pi_{k+1,-k}$ is the automorphic induction of the character $z \mapsto z^{-k}$.

When $k \geq 2$, $\pi_{k,s}$ is in the *discrete series*. When $k = 1$ $\pi_{k,s}$ is a *limit of discrete series* representation. The distinction of terminology reflects the fact that the Eichler-Shimura isomorphism relating cohomology of modular curves with holomorphic modular forms does not work for forms of weight 1.

1.3.3. Others.

All other irreducible admissible representations of $GL(2, \mathbb{R})$ are either finite-dimensional (and are classified by standard Lie theory) or irreducible principal series, constructed just as in the non-archimedean case. There is no analogue of supercuspidal representations for real reductive groups.

1.4. Classification of irreducible representations of $D^\times$.

In what follows, F is a local field, and D is the unique non-split quaternion algebra over F. If $F = \mathbb{R}$, D is the algebra of Hamiltonian quaternions; otherwise, D is the four-dimensional division algebra over F that corresponds to the element $\frac{1}{2} \in \mathbb{Q}/\mathbb{Z} \xrightarrow{\sim} Br(F)$. The multiplicative group $D^\times$, which we denote J, is compact modulo its center $Z \xrightarrow{\sim} F^\times$. It follows from Schur's lemma and the representation theory of compact groups that any admissible irreducible representation of J is necessarily finite-dimensional. We will not dwell on the construction of representations of J but will present the Jacquet-Langlands correspondence.

Let $N_D : J \to F^\times$ denote the reduced norm. We can think of J as the group of F-points of an algebraic group $\mathbb{J}$ over F. Then $\mathbb{J}$ is a twisted inner form of $GL(2)$, and in particular there is an isomorphism

$$\mathbb{J}(\bar{F}) \xrightarrow{\sim} GL(2, \bar{F}),$$

well-defined up to conjugation. Thus any element of $J \subset \mathbb{J}(\bar{F})$ can be considered a 2×2 invertible matrix with coefficients in $\bar{F}$; in particular it has a characteristic polynomial, independent of the choice of isomorphism above, and with coefficients in F. We say $j \in J$ and $g \in G$ are *associated* if they have the same characteristic polynomial; this implies g is elliptic (including non-regular elliptic elements, necessarily in the center of G).

When $F = \mathbb{R}$, the derived subgroup $J^{der} \subset J$ is isomorphic to the compact Lie group $SU(2)$. Any irreducible representation τ of J is determined by its restriction to $SU(2)$ and by its central character, and indeed by the character by which $Z^0 = \mathbb{R}^{\times,+} \subset Z$ acts on τ. An irreducible representation of $SU(2)$ is in turn determined by its dimension. For $k > 1$ in $\mathbb{Z}$ and $s \in \mathbb{C}$, let $\tau_{k,s}$ be the representation of dimension $k-1$ on which $t \in Z^0$ acts by t^s. Then obviously

Proposition 1.4.1. *The map $\tau_{k,s} \leftrightarrow \pi_{k,s}$ is a bijection between the (equivalence classes of) irreducible representations of J and the discrete series of $GL(2,\mathbb{R})$.*

The bijection above is called the *Jacquet-Langlands correspondence* over $\mathbb{R}$. Note that $\pi_{1,s}$ is excluded.

Now assume F is nonarchimedean, and let $\mathcal{A}^D$ denote the set of equivalence classes of irreducible admissible representations of J. Let $\mathcal{A}^{ds} \subset \mathcal{A}_{2,F}$ be the set of discrete series representations. The following theorem is considerably deeper than Proposition 1.4.1.

Theorem 1.4.2. *There is a natural bijection (Jacquet-Langlands correspondence)*

$$JL : \mathcal{A}^D \to \mathcal{A}^{ds}.$$

This bijection has the following natural properties

(i) *JL preserves local L and ε factors in the sense of Godement-Jacquet;*
(ii) *Let χ be a continuous character of $F^\times$ and let $\tau_\chi = \chi \circ N_D$. Then $JL(\tau_\chi) = St(\chi)$.*
(iii) *Let $j \in J$ and $g \in G$ be associated and assume g is regular. Then for any $\tau \in \mathcal{A}^D$,*

$$\chi_{JL(\tau)}(g) = -Tr(\tau(j)).$$

The last identity is proved by means of the Selberg trace formula.

Proposition 1.4.1 is exactly the real version of Theorem 1.4.2.

2. The global Jacquet-Langlands correspondence

In the present section, F is a totally real number field, and D/F is a quaternion algebra; we let $G = GL(2)_F$, $J = GL(1,D)$, viewed as algebraic groups over F. Automorphic representations of G and J have been defined in an earlier talk at Luminy; However, this material is not available, so I include a brief discussion here of automorphic representations of $GL(2)$, and implicitly of $GL(n)$.

A holomorphic modular form F of weight k is a holomorphic function on the upper half-plane $\mathfrak{H}$ satisfying a certain symmetry. The group $SL(2,\mathbb{R})$ acts by linear-fractional transformations on $\mathfrak{H}$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}(z) = \frac{az+b}{cz+d}$$

and F is assumed to transform according to a certain rule, depending on k, for $\gamma \in \Gamma$ where $\Gamma \subset SL(2,\mathbb{Q}) \subset SL(2,\mathbb{R})$ is a congruence subgroup. There is also a growth condition near the boundary of $\mathfrak{H}$. I assume this is all familiar and I mention this in order to motivate the definition of automorphic forms in general. With little effort one extends F to a function on $\mathfrak{H} \coprod \bar{\mathfrak{H}}$ (upper + lower halfplanes), which is identified with $GL(2,\mathbb{R})/SO(2) \cdot \mathbb{R}^\times$. A standard lifting procedure lifts F to a C^∞ function f on $\Gamma\backslash GL(2,\mathbb{R})$, and thence to a function

$$f : GL(2,\mathbb{Q})\backslash GL(2,\mathbf{A}) \to \mathbb{C}$$

where in each case the transformation rule itself undergoes a transformation. The original holomorphy condition is inherited by f in another form: the C^∞ function f is a solution to a differential equation corresponding to the Cauchy-Riemann equations on $\mathfrak{H}$. For purposes of generalization, it is best to think of a holomorphic form of weight 2 as a holomorphic 1-form on the open complex curve $\Gamma\backslash\mathfrak{H}$, and then just as a cohomology class, possibly with some funny behavior at the missing points. Forms of weight $k > 2$ define cohomology with twisted coefficients by the Eichler-Shimura isomorphism. That is what generalizes for $n > 2$.

Following earlier work of Mordell, Hecke discovered the operators that bear his name on the spaces of modular forms, and more importantly, that the simultaneous eigenfunctions for all these operators could be assigned Dirichlet series with analytic continuation and functional equations. The extra adelic variables allow a natural definition of the Hecke operators. A (discrete) *automorphic representation* of $GL(2,\mathbb{Q})$ is a direct summand of

$$L^2(GL(2,\mathbb{Q})\backslash GL(2,\mathbf{A})/\mathbb{R}^\times)$$

which defines an irreducible representation π^* of $GL(2,\mathbf{A})$ under the right regular representation. When one applies the lifting procedure of the previous paragraph to one of the eigenforms Hecke considered, its right $GL(2,\mathbf{A})$-translates form an automorphic representation in this sense, and all the eigenvalues of the Hecke operators can be recovered from the structure of the abstract representation. Most discrete automorphic representations are *cuspidal*, which means they decrease rapidly near infinity, and we only consider cuspidal representations.

It is better to take the subspace π^1 of π^* consisting of functions that are C^∞ in the variables in $GL(2,\mathbb{R})$ and locally constant in the non-archimedean variables. A vector $v \in \pi^1$ is *K-finite* if the space generated by the translates of v by K is finite-dimensional. Inside π^1 the subspace π of K-finite vectors forms a Harish-Chandra module, or rather is the tensor product of a Harish-Chandra module with an irreducible representation of $GL(2,\mathbf{A}^f)$. It is this π that we call a (cuspidal) automorphic representation. It is thus an irreducible object in the category of $(\mathfrak{g},K) \times GL(2,\mathbf{A}^f)$-modules.

The same definition holds when $\mathbb{Q}$ is replaced by any number field F and when 2 is replaced by any n. We let $G = GL(n)_F$. I will describe the theory of representations of $G(F_v)$ below. For v non-archimedean it is roughly analogous to the case already considered of $GL(2)$. On the other hand, for v archimedean there is no natural collection of $(\mathfrak{g}_\mathbb{C},K)$-modules comparable to the holomorphic representations.[1] Here $K = \prod_{v|\infty} K_v$ is a maximal compact subgroup of $G_\infty = \prod_{v|\infty} G(F_v)$,

[1] Indeed, there is no theory of holomorphic modular forms related to automorphic representations of $GL(n)$ for $n > 2$; one gets around this problem by working with unitary groups, as we will see later.

so K_v is either isomorphic to $\mathbb{R}^\times \cdot O(n)$ (v real) or $\mathbb{C}^\times \cdot U(n)$, ($v$ complex). If π is an (irreducible cuspidal) automorphic representation of G, then it factors

$$\pi = \otimes'_v \pi_v$$

(restricted tensor product over places v of F), where for each v, π_v is an admissible irreducible representation of $G(F_v)$ in the sense described in §1. Likewise for J. The function $L(s,\pi) = \prod_{v \nmid \infty} L(s,\pi_v)$ converges in a right half plane and extends to an entire function with functional equation (Godement-Jacquet).

I return to the case $n = 2$. Let $\Sigma(D)$ be the set of places of F where D ramifies, i.e. where $D_v \not\simeq GL(2,F_v)$. Recall that $\Sigma(D)$ has an even number of elements. Conversely, if Σ is any finite set of even cardinality, there is a quaternion algebra D, unique up to isomorphism, such that $\Sigma = \Sigma(D)$.

Theorem 2.1 (Jacquet-Langlands correspondence). *Let $\tau = \otimes'_v \tau_v$ be a cuspidal automorphic representation of J. For $v \notin \Sigma(D)$, τ_v is an irreducible admissible representation of $GL(2,F_v)$, which we denote π_v; its equivalence class is well-defined. For $v \in \Sigma(D)$, set $\pi_v = JL(\tau_v)$. Let $\pi = \otimes'_v \pi_v$. Then π is isomorphic to a cuspidal automorphic representation of G, and we write $\pi = JL(\tau)$.*

Conversely, let π be a cuspidal automorphic representation of G. Suppose that for all $v \in \Sigma(D)$, $\pi_v \in \mathcal{A}^{ds}$. In particular, for each $v \in \Sigma(D)$, there is $\tau_v \in \mathcal{A}^{D_v}$ such that $\pi_v = JL(\tau_v)$. For $v \notin \Sigma(D)$ we set $\tau_v = \pi_v$ as above. Then $\tau = \otimes'_v \tau_v$ is isomorphic to a cuspidal automorphic representation of J, and $\pi = JL(\tau)$. Moreover, τ has multiplicity one in the space of all automorphic forms on J.

The bijection JL identifies $L(s,\pi) = L(s,\tau)$ as well as the functional equations.

In other words, JL defines a bijection between the cuspidal automorphic representations of J and the set of cuspidal automorphic representations π of G such that π_v is in the discrete series for all $v \in \Sigma(D)$.

The proof of the Jacquet-Langlands correspondence is now considered the simplest non-trivial application of the Selberg trace formula to Langlands functoriality. It is only directly relevant when $n = 2$ but is the model for the functoriality between inner forms of unitary groups to be discussed below.

2.2. Remark. The Jacquet-Langlands correspondence was constructed in many cases by Shimizu, using the Weil representation, which concretely means theta functions for orthogonal groups in four variables. Following the work of Jacquet-Langlands, Shimizu showed that the Jacquet-Langlands correspondence can be obtained in all cases by means of theta functions. However, proof of the properties of the bijection seems to require the trace formula.

Suppose $\Sigma(D)$ contains all but one of the archimedean places of F. Then J is the algebraic group associated to a Shimura curve. Most of the global Galois representations discussed in the following section are realized in the ℓ-adic cohomology of such Shimura curves with twisted coefficients.

3. Associated global Galois representations

Notation is as in the previous section. Henceforward we will only consider cuspidal automorphic representations π of $GL(2,F)$ such that π_v is a holomorphic representation, i.e. a $\pi_{k,s}$, for all archimedean v. Such π are (naturally) called *holomorphic*. To a holomorphic π we can associate a number field $E(\pi)$ with the property that $\pi_f := \otimes'_{v \nmid \infty} \pi_v$ has an $E(\pi)$-rational model as abstract representation. If π is associated to the elliptic modular normalized newform f, then $E(\pi) = E(f)$ is the extension of $\mathbb{Q}$ generated by the Fourier coefficients of f.

If v is a finite place of F, let G_v denote a decomposition group in $Gal(\overline{\mathbb{Q}}/F)$.

Theorem (Eichler, Shimura, Deligne, Langlands, Carayol, Scholl, T. Saito,??; Deligne-Serre, Rogawski, Tunnell, Ohta). *Assume $\pi_v = \pi_{k,2-k}$ for all archimedean v. There is a compatible system of λ-adic representations (of geometric type), as λ runs through non-archimedean completions of $E(\pi)$:*

$$\rho_{\lambda,\pi} : Gal(\overline{\mathbb{Q}}/F) \to GL(2, E(\pi)_\lambda)$$

that is associated to π in the sense that, for every ℓ and every place v of F prime to the residue characteristic ℓ of λ,

$$\rho_{\lambda,\pi} \mid_{G_v} = \mathcal{L}^*(\pi_v) \tag{1}$$

where $\mathcal{L}^*$ is the (motivically normalized) local Langlands correspondence.

Moreover, let v be a prime of F dividing ℓ. Then the representation $\rho_{\lambda,\pi} \mid_{G_v}$ is potentially semistable, the Hodge-Tate numbers of $\rho_{\lambda,\pi}$ at v are 0 and $k-1$, with multiplicity one, and

$$D(\rho_{\lambda,\pi} \mid_{G_v}) = \mathcal{L}^*(\pi_v) \tag{2}$$

where D is the Fontaine functor associating to any potentially semistable representation of G_v a representation of the Weil-Deligne group of F_v.

For the final step for Hilbert modular forms, there still seem to be some missing cases.

The motivically normalized local Langlands correspondence involves a half-integral shift, so that

$$L(s, \mathcal{L}(\pi_v)) = L(s - \frac{1}{2}, \pi_v). \tag{3}$$

This guarantees that the λ-adic representations are realized over completions of a fixed $E(\pi)$ and not over the (infinite) extension of $E(\pi)$ obtained by adjoining square roots of ℓ for all ℓ.

A representation of the form $\rho_{\lambda,\pi}$, or its reduction $\bar{\rho}_{\lambda,\pi}$ to characteristic ℓ, is called *modular*, or sometimes *automorphic*. The goal of modularity theorems is to start with a Galois representation ρ_λ or $\bar{\rho}_\lambda$ and to show that it arises from some π by the construction in the theorem.

When $k = 1$, $\rho_{\lambda,\pi}$ is an Artin representation, whose construction is by the method of Deligne-Serre.

4. Base change and global automorphic induction

4.1 Base change.

Now suppose F'/F is a cyclic Galois extension of totally real fields, of prime degree. Let $G' = GL(2)_{F'}$. Let $\pi \simeq \otimes'_v \pi_v$ be a cuspidal automorphic representation of G, which we assume holomorphic of parallel weight k (though this is irrelevant). We construct a holomorphic automorphic representation $\pi' \simeq \otimes'_w \pi'_w$ of G' by base change, where w runs over places of F', as follows:

(a) Suppose v is a place of F that splits (completely) in F'. Then for any divisor w of v, $G'(F'_w) \simeq G(F_v)$. We let $\pi'_w = \pi_v$. This includes all real places.

(b) Suppose v does not split, and let w be the unique place of F' dividing v. We let $\pi'_w = BC_{F'_w/F_v}(\pi_v)$, defined as above.

Theorem 4.1.1 (Saito-Shintani, Langlands). *(a) The representation* $\pi' = BC_{F'/F}(\pi)$ *is an automorphic representation of* G'. *If* σ *is a generator of* $Gal(F'/F)$, *then* $\pi' \circ \sigma \xrightarrow{\sim} \pi'$, *where* $Gal(F'/F)$ *acts on the equivalence class of* π' *by acting on the underlying adele group.*

(b) Conversely, suppose π' *is a cuspidal automorphic representation of* G' *such that* $\pi' \circ \sigma \xrightarrow{\sim} \pi'$. *Then* π' *is of the form* $BC_{F'/F}(\pi)$ *for some cuspidal automorphic representation* π *of* G.

(c) Suppose $[F' : F] = 2$ *and let* η *be the quadratic character of the idèles of* F *associated to the extension* F'. *Suppose* $\pi \otimes \eta \circ \det \simeq \pi$ *as abstract representations. Then there is a Hecke character* χ *of the idèles of* F' *such that* π *is obtained from* χ *by automorphic induction (see §4.4 below). In particular,* $k = 1$.

(d) Suppose we are not in case (b). Then π' *is a cuspidal automorphic representation of* G' *that is holomorphic of parallel weight* k. *Moreover,* $E(\pi') = E(\pi)$, *and for every finite place* λ *of* $E(\pi)$,

$$\rho_{\lambda,\pi'} = \rho_{\lambda,\pi} \mid_{Gal(\overline{\mathbb{Q}}/F')} .$$

The fact that (c) can only arise when $k = 1$ can be seen by analyzing the condition $\pi \otimes \eta \circ \det \simeq \pi$ locally at archimedean primes.

By induction, one can construct $BC_{F'/F}(\pi)$ for any solvable extension F'/F. The condition that the base change remain cuspidal is more subtle in general but can be deduced from Theorem 4.1.1. It is clear that the equivalence class of $BC_{F'/F}(\pi)$ is invariant under the natural action of $Gal(F'/F)$. If $BC_{F'/F}(\pi)$ is cuspidal, it follows from (d) and the characterization of $\rho_{\lambda,\pi}$ that $\rho_{\lambda,BC_{F'/F}(\pi)}$ is invariant under $Gal(F'/F)$. Conversely,

Theorem 4.1.2. *Let* F'/F *be a solvable extension, and let* π' *be a cuspidal automorphic representation of* $G' = GL(2)_{F'}$ *of parallel weight* $k \geq 2$. *Suppose* $\rho_{\lambda,\pi'}$ *extends to a representation of* $Gal(\overline{\mathbb{Q}}/F)$. *Then there is a cuspidal automorphic representation* π *of* $G = GL(2)_F$, *of parallel weight* k, *such that* $\pi' = BC_{F'/F}(\pi)$.

In other words, if ρ is a two-dimensional λ-adic representation of $Gal(\overline{\mathbb{Q}}/F)$ that becomes modular over a solvable totally real extension, then ρ is already modular over F. I think as stated in Theorem 4.1.2, the condition that $k \neq 1$ is superfluous. However, it is **definitely not** known in general that an Artin representation that becomes modular over a solvable extension is necessarily modular; otherwise the Artin conjecture would be known for all solvable Galois representations. Quite a few famous mathematicians have allowed themselves to be tricked on this point.

The problem is that the characterization of π such that $BC_{F'/F}(\pi) = \pi'$, in terms of π', is given by a local relation of characters between π_v and π'_v for all v. Global descent is determined up to twist by a global character of the cyclic extension, but locally there is a separate local character (and usually two) at each non-split place in F'/F. This has the following consequence. Suppose $F' \supset F \supset E$ is a triple of number fields, with F'/E Galois and solvable, and F'/F and F/E both cyclic of prime order. Let π' be a cuspidal automorphic representation of $GL(2, F')$ that is invariant under $Gal(F'/E)$. Then π' descends to a cuspidal automorphic representation π of $GL(2, F)$. Now if $\sigma \in Gal(F/E)$, the most we know is that

$$\pi_v \circ \sigma \simeq \pi_v \otimes \alpha_v \circ \det$$

for all primes v, where α_v is a local Galois character, and this does not suffice for descent; indeed, one doesn't even know *a priori* that the α_v fit together to form a global character of $Gal(F/E)$.

However, if one already knows that π is associated to a global (irreducible) Galois representation – because π_∞ is a holomorphic representation, for example – then it is easy to see that π can be chosen invariant under $Gal(F/E)$, so that descent can continue. (This simple observation may have been applied for the first time in my proof of the local Langlands conjecture in the tame case.)

4.2 Global automorphic induction of Hecke characters, imaginary quadratic case.

The following construction works with $\mathbb{Q}$ replaced by any base field. Let $E/\mathbb{Q}$ be a quadratic field, and let $\sigma \in Gal(E/\mathbb{Q})$ be the non-trivial automorphism. Let $\chi : E_{\mathbf{A}}^\times/E^\times \to \mathbb{C}^\times$ be a Hecke character. Define the representation $\pi(\chi) = \otimes_v \pi_v(\chi_v)$ as follows:

(a) Suppose v is a place of $\mathbb{Q}$ that splits in E, $v = w \cdot w^\sigma$. Let $\pi_v(\chi_v)$ be the principal series $I(\chi_w, \chi_{w^c})$.

(b) Suppose v does not split, and let w be the unique place of E dividing v. We let $\pi_v = AI_{E_w/\mathbb{Q}_v}(\chi_w)$, defined as above.

Theorem 4.2.1 (Hecke, Maass, Jacquet-Langlands). *(a) The representation* $AI_{E/\mathbb{Q}}(\chi) = \pi(\chi)$ *is an automorphic representation of* G'*. If* η_E *is the quadratic character of the idèles of* $\mathbb{Q}$ *associated to the extension* E*, then*

$$\pi(\chi) \otimes \eta_E \circ \det \simeq \pi(\chi).$$

(b) Conversely, suppose π *is a cuspidal automorphic representation of* $GL(2)_\mathbb{Q}$ *such that* $\pi(\chi) \otimes \eta_E \circ \det \simeq \pi(\chi)$*. Then* π *is of the form* $\pi(\chi)$ *for some Hecke character* χ *of* E*.*

(c) The automorphic representation $\pi(\chi)$ *is cuspidal if and only if* $\chi \neq \chi \circ \sigma$*, or equivalently if and only if* χ *does not factor through the norm down to the idèles of* $\mathbb{Q}$*.*

As indicated, these automorphic representations, or rather their corresponding modular forms (resp. Maass forms) were constructed by other means by Hecke (resp. Maass). They were also constructed by the converse theorem in various cases by Weil, Shimura, and Jacquet-Langlands.

We consider the imaginary and real quadratic cases separately.

Theorem 4.2.2. *Suppose* E *is imaginary quadratic, and suppose* $\chi_\infty(z) = z^{1-k}$ *for some* $k > 0$*. (In particular,* χ *never factors through the norm if* $k > 1$*.) Then* $\pi(\chi)$ *is holomorphic of weight* k*, and is always cuspidal if* $k > 1$*. Moreover,* $E(\pi(\chi))$ *is contained in the number field* $E(\chi)$ *generated by the values of* χ *on the finite idèles, and for all places* λ *of* $E(\chi)$*,* $\rho_{\lambda,\pi(\chi)}$ *is induced to* $Gal(\overline{\mathbb{Q}}/\mathbb{Q})$ *from the* λ*-adic character* $Gal(\overline{\mathbb{Q}}/E) \to E(\chi)_\lambda^\times$*.*

Theorem 4.2.3. *Suppose* E *is real quadratic, and let* v, v' *be the two real places of* E*. Suppose* χ_v *is the trivial character and* χ'_v *is the sign character; in particular,* $\chi \neq \chi \circ \sigma$*. Then* $\pi(\chi)$ *is cuspidal holomorphic of weight* 1*. Moreover, for all* λ*,* $\rho_{\lambda,\pi(\chi)}$ *is an Artin representation induced from the (finite) Galois character of* $Gal(\overline{\mathbb{Q}}/E)$ *associated to* χ*.*

Theorems 4.2.2 and 4.2.3, together with the theorem of Serre and Henniart on locally algebraic ℓ-adic representations, imply that every two-dimensional λ-adic representation of $Gal(\overline{\mathbb{Q}}/\mathbb{Q})$ that is Hodge-Tate and monomial (induced from a one-dimensional character of the Galois group of a quadratic extension) is modular.