ARITHMETIC APPLICATIONS OF THE LANGLANDS PROGRAM

Michael Harris

UFR de Mathématiques
Université Paris 7
2 Pl. Jussieu
75251 Paris cedex 05, FRANCE

Introduction

The Galois group was introduced as a means of studying roots of polynomials, but it now more often finds itself an object of study in its own right. The method of choice for studying a group Γ is by describing its representations

$$\rho : \Gamma \to GL(V)$$

where V is a vector space or, more generally, a module over a commutative ring. When Γ is a Galois group it is a compact, indeed profinite, topological group, so one restricts one's attention to ρ that are *continuous* with respect to a given topology on V.

Suppose for the moment that $\Gamma = \Gamma_F = Gal(\bar{F}/F)$ for a number field F, a finite extension of $\mathbb{Q}$ with algebraic closure $\bar{F} \simeq \overline{\mathbb{Q}}$, and V is a complex vector space; then the image of a continuous ρ is necessarily finite, and its kernel defines a finite Galois extension L/F; in other words ρ factors through $Gal(L/F)$. One can't get very far in determining the representations of $Gal(L/F)$ as long as its elements remain anonymous. Since we are doing number theory, we can label the elements of $Gal(L/F)$ by *primes*, as follows. Assume for the moment that $F = \mathbb{Q}$, and let $\mathcal{O} = \mathcal{O}_L$ denote the integer ring of L. Let p be a prime number, and consider a ring homomorphism $v : \mathcal{O} \to \overline{\mathbb{F}}_p$. The Galois group $Gal(\overline{\mathbb{F}}_p/\mathbb{F}_p)$ is topologically generated by the *Frobenius element* σ_p with the property that

$$\sigma_p(x) = x^p, \ \forall x \in \overline{\mathbb{F}}_p.$$

It is important and not difficult to prove that there exists an element $\sigma_v \in Gal(L/\mathbb{Q})$ with the property that, for all $a \in \mathcal{O}$,

$$v(\sigma_v(a)) = \sigma_p(v(a)). \tag{0.1}$$

Institut des Mathématiques de Jussieu, U.M.R. 7586 du CNRS. Membre, Institut Universitaire de France.

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

One can say roughly that $\sigma_v(a)$ is "congruent to a^p" modulo p, except that there is generally more than one way to "reduce an element of $\mathcal{O}$ modulo p"; the choice of ring homomorphism v above precisely corresponds to the different ways of doing so. More precisely still, the different v's as above correspond to the different prime ideals of $\mathcal{O}$ dividing p.

It is not obvious, but still not difficult, to show that for all but finitely many primes p there is exactly one element $\sigma_v \in Gal(L/\mathbb{Q})$ satisfying (0.1). The finitely many exceptional primes are said to *ramify* in $L/\mathbb{Q}$; they are the source of much of the excitement of number theory, but we ignore them for the time being. Assuming p is unramified, we let $Frob_v \in Gal(L/\mathbb{Q})$ denote the element σ_v^{-1}. This labels elements of $Gal(L/\mathbb{Q})$ by prime ideals of L, but if v and v' are two homomorphisms as above, for the same prime number p, then they are conjugate in $Gal(L/\mathbb{Q})$. Thus conjugacy classes in $Gal(L/\mathbb{Q})$ can be labeled by prime numbers, and this is consistent if L is replaced by another field L' such that ρ factors through $Gal(L'/\mathbb{Q})$. Moreover, each conjugacy class in $Gal(L/\mathbb{Q})$ is labeled by infinitely many primes: indeed the *Chebotarev density theorem*, a tool of fundamental importance in modern number theory, asserts that the density of primes labeling a given conjugacy class $\mathcal{C}$ is proportional to the number of elements of $\mathcal{C}$. For general F, the conjugacy classes in $Gal(L/F)$ are labeled in a similar way by prime ideals of $\mathcal{O}_F$, and the Chebotarev density theorem remains valid.

In an abelian $Gal(L/F)$, conjugacy classes are simply elements. When $F = \mathbb{Q}$, the Chebotarev density theorem for abelian extensions amounts to Dirichlet's theorem on primes in an arithmetic progression; the prime numbers that label a given element of $Gal(L/\mathbb{Q})$ are those that satisfy an appropriate congruence condition. *Class field theory* asserts that this is still true for general F, using a roundabout definition of congruence applicable to prime ideals, and that the congruence condition determines the abelian extension L. The theory is completed by the *existence theorem*, due to Teiji Takagi, which asserts, roughly, that, conversely, every congruence condition of the appropriate type corresponds to an abelian extension L/F.

It is a great honor to be asked to deliver one of the Takagi lectures in Tokyo, on the occasion of the fiftieth anniversary of the death of the individual considered by many to be the founder of modern mathematics in Japan. But this honor comes in the present case with a special responsibility, since my task is to provide an account of recent developments in the non-abelian generalization of the class field theory in the creation of which Teiji Takagi played a central role. Takagi's existence theorem has the following consequence in modern language. Let $\mathbf{A}_F^\times$ denote the idèle group of F. Then there is a map from the set $\mathcal{A}_1(F)_{fin}$ of continuous homomorphisms

$$\chi : \mathbf{A}_F^\times/F^\times \to \mathbb{C}^\times \tag{0.2}$$

with finite image, to the set $\mathcal{G}_1(F)_{fin}$ of continuous characters

$$\rho : \Gamma_F \to \mathbb{C}^\times. \tag{0.3}$$

Combining Takagi's existence theorem with the other theorems of class field theory, one finds that this map is in fact a *bijection*. The "congruence condition" to which I alluded above is an open subgroup of finite index in the idèle class group $\mathbf{A}_F^\times/F^\times$, for example the kernel of χ, which can be determined explicitly in terms of the fixed field of the kernel of ρ.

One can view the idéle class group as a means for encoding simultaneously all the congruence conditions that arise in class field theory. What replaces the congruence condition when the extension L/F is not abelian? From Langlands' perspective, it is more natural to take as starting point the representation ρ of dimension $n = \dim V$, rather than the Galois extension L/F itself. To express the appropriate congruence condition, one then needs the entire representation theory of the group $GL(n)$ over the adèles of F. Thus non-abelian class field theory becomes a relation between harmonic analysis on Lie groups and Galois theory.

Although the possibility of a non-abelian version of class field theory was prefigured in work of Taniyama and Shimura, it first appeared as a full-fledged research program ten years after Takagi's death in Langlands 1970 article [L70]. In what is perhaps the most familiar special case of Langlands' functoriality conjectures, the set $\mathcal{G}_1(F)_{fin}$ of (0.2) is replaced by the set $\mathcal{G}_n(F)_{fin}$ of equivalence classes of irreducible n-dimensional continuous complex representations of Γ_F, and $\mathcal{A}_1(F)_{fin}$ is replaced by a certain subset $\mathcal{A}_n(F)_{fin}$ of the set of cuspidal automorphic representations of $GL(n, \mathbf{A}_F)$. This special case implies the Artin conjecture for F, and for this reason naturally attracted the attention of number theorists.[1] Langlands full functoriality conjecture asserts that Galois-theoretic data, broadly construed, predicts relations between automorphic representations of different connected reductive groups.

The functoriality conjecture is at the heart of the Langlands program and will undoubtedly remain as a challenge to number theorists for many decades to come. Shortly after formulating his program, however, Langlands proposed to test it in two interdependent settings. The first was the framework of *Shimura varieties*, already understood by Shimura as a natural setting for a non-abelian generalization of the Shimura-Taniyama theory of complex multiplication. The second was the phenomenon of endoscopy, which can be seen alternatively as a classification of the *obstacles* to the stabilization of the trace formula or as an *opportunity* to prove the functoriality conjecture in some of the most interesting cases. After three decades of research, much of it by Langlands and his associates, these two closely related experiments are coming to a successful close, at least for classical groups, thanks in large part to the recent proof of the so-called *Fundamental Lemma* by Waldspurger, Laumon, and especially Ngô.

My primary interest in these lectures is to give an account of these developments insofar as they are relevant to the Galois groups of number fields. While the Langlands program itself is now undertaking a new series of experiments, the most ambitious of which has been proposed by Langlands himself, algebraic number theorists are just beginning to take stock of the new information provided by the successful resolution of the problem of endoscopy and the analysis of the most important classes of Shimura varieties. I will devote special attention to the application in this setting of the methods pioneered by Wiles, and developed further by Taylor, Kisin, and others, in his proof of statements in non-abelian class field theory that imply Fermat's Last Theorem.

In preparing this lecture, I have greatly benefited from the advice of a number of colleagues, including J. Arthur, L. Clozel, M. Emerton, M. Kisin, J.-P. Labesse, C. Moeglin, B. C. Ngô, D. Ramakrishnan, R. Taylor, and J.-L. Waldspurger. I am

[1] The important recent progress on this special case when $n = 2$ and $F = \mathbb{Q}$ will be reviewed in Part III below.

grateful to the referees for their very careful reading of the manuscript, and for their numerous suggestions for clarification.

Part I. Review of Langlands' functoriality conjectures

I.1. Formalism of the L-group

We let F be either a local field or a number field, and let Γ denote either $Gal(\bar{F}/F)$ or the Weil group $W_{\bar{F}/F}$. Let G be a connected reductive group over F. Langlands' theory of the L-group begins with the observation that a connected reductive algebraic group $\mathcal{G}$ over an algebraically closed field (of characteristic zero, for simplicity) can be reconstructed up to canonical isomorphism in terms of the *based root datum* $\Psi_0(\mathcal{G}) = (X^*, \Delta^*, X_*, \Delta_*)$. Here T is a maximal torus of $\mathcal{G}$, $X^* = X^*(T)$, resp. X_* is its group of characters (resp. cocharacters), $B \supset T$ is a Borel subgroup, Δ^* is the set of positive simple roots of T in B, Δ_* the set of positive simple coroots. These data depend on a number of choices, but for any two choices the corresponding data are *canonically* isomorphic, hence the notation $\Psi_0(\mathcal{G})$ is justified.

The quadruple $(X^*, \Delta^*, X_*, \Delta_*)$ satisfies a collection of axioms including as a subset the axioms for a root system satisfied by the first two items. It suffices to mention that any quadruple Ψ of the same type, with X^* and X_* finitely generated free abelian groups, and $\Delta^* \subset X^*$ and $\Delta_* \subset X_*$, that satisfy these axioms comes from a connected reductive group. The dual based root datum $\hat{\Psi} = (X_*, \Delta_*, X^*, \Delta^*)$, obtained by switching the first two items with the last two, also satisfies these axioms, and therefore defines up to canonical isomorphism a connected reductive algebraic group $\hat{\mathcal{G}}$ over $\mathbb{C}$ with $\Psi_0(\hat{\mathcal{G}}) \xrightarrow{\sim} \hat{\Psi}$.

We write $\hat{G}$ for $\hat{G}_{\bar{F}}$. The F-rational structure on $G_{\bar{F}}$ translates to an action of Γ on $\Psi_0(G_{\bar{F}})$, preserving the natural pairing between characters and cocharacters, and thus, almost, to an action of Γ on $\hat{G}$. To define such an action unambiguously one needs to choose a *splitting* ("épinglage") of $\hat{G}$, i.e., a triple $(\hat{T}, \hat{B}, \{X_\alpha, \alpha \in \hat{\Delta}\})$, where $\hat{T}$ is a maximal torus of $\hat{G}$, $\hat{B} \supset \hat{T}$ a Borel subgroup, $\hat{\Delta}$ the set of positive simple roots of $\hat{T}$ in $\hat{B}$, and X_α is a non-zero element of the root space $\hat{\mathfrak{g}}_\alpha \subset Lie(\hat{B})$ for every $\alpha \in \hat{\Delta}$. An action of Γ on $\hat{G}$ that fixes some splitting is called an L-action, and an L-group ${}^LG$ of G is the semi-direct product of $\hat{G}$ with Γ with respect to a fixed L-action.

Concretely, we will need the following L-groups in the discussion below:

$$G_n = GL(n,F),\ {}^LG_n = GL(n,\mathbb{C}) \times \Gamma.$$

$$H_n = SO(2n+1,F),\ {}^LH_n = Sp(2n,\mathbb{C}) \times \Gamma.$$

$$J_n = Sp(2n,F),\ {}^LJ_n = SO(2n+1,\mathbb{C}) \times \Gamma.$$

Let F'/F be a quadratic extension and V an n-dimensional hermitian vector space over F'.

$$U_n = U(V),\ {}^LU_n = GL(n,\mathbb{C}) \rtimes \Gamma.$$

Letting $c \in Gal(F'/F)$ denote the non-trivial element, the semidirect product is determined by the non-trivial action of c on $\hat{G}$. Let Φ_n be the matrix whose ij

entry is $(-1)^{i+1}\delta_{i,n-j+1}$:

$$\Phi_n = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & -1 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ (-1)^{n-1} & 0 & \dots & 0 & 0 \end{pmatrix} \tag{I.1.1}$$

Then let c act on $\hat{G} = GL(n,\mathbb{C})$ as

$$c(g) = \Phi_n {}^t g^{-1} \Phi_n^{-1}. \tag{I.1.2}$$

(The matrix Φ_n is introduced in order to preserve the standard épinglage.)

An *L-homomorphism* is a map of L-groups that commutes with the projection to Γ. There are obvious L-homomorphisms

$$^L H_n \to {}^L G_{2n};\ {}^L J_n \to {}^L G_{2n+1}.$$

When F is local, Langlands' functoriality conjectures predict that these give rise to *functorial transfer* of irreducible admissible representations of $SO(2n+1,F)$ to irreducible admissible representations of $GL(2n,F)$ (respectively from $Sp(2n,F)$ to $GL(2n+1,F)$). There is a slightly less obvious L-homomorphism

$$^L U_n \to {}^L R_{F'/F} GL(n)_{F'}$$

(cf. [Mi, 4.1.2]) that should give rise to transfer, in this case also called *base change*, of irreducible admissible representations of $U_n(F)$ to $GL(n,F')$; this is discussed in detail in §II.5. The transfer map is in general not injective, and the fibers above representations of the target group are called *L-packets*. When F is global, transfer associated to an L-homomorphism ${}^L G' \to {}^L G_n$ is supposed to take (packets of) automorphic representations of G' to automorphic representations of G_n.

The concrete *description* of the functorial transfer map is given in terms of (local and global) Langlands parameters, to which we turn presently. The primary *method* for proving functorial transfer is provided by the comparison of stable trace formulas. That the comparison of trace formulas actually yields the conjectured description of functorial transfer is the burden of a collection of results about harmonic analysis on (real and p-adic) Lie groups, of which the most important is the *Fundamental Lemma*. Now that the Fundamental Lemma is a theorem, the instances of functorial transfer listed above are now either proved or in the process of being proved, along with some other cases we have omitted (for example, the case of even orthogonal groups).

I.2. Local Langlands parameters

In this section F is a local field. A *local Langlands parameter* for the group G is an equivalence class of homomorphisms of the appropriate type from the Weil-Deligne group of F to ${}^L G$. These parameters look very different for archimedean and non-archimedean local fields, and there is more than one way to define Weil-Deligne parameters when F is non-archimedean. We consider these issues in turn.

For F archimedean, the Weil-Deligne group is just the Weil group W_F. When $F = \mathbb{C}$, $W_F = \mathbb{C}^\times$, ${}^L G = \hat{G}$, and a local Langlands parameter is then just an equivalence class of homomorphisms

$$\phi : W_F = \mathbb{C}^\times \to \hat{G}$$

whose image is contained in a torus (so $z \mapsto \begin{pmatrix} 1 & \log|z| \\ 0 & 1 \end{pmatrix}$ is excluded); two homomorphisms ϕ and ϕ' are equivalent if they are conjugate. For $F = \mathbb{R}$, W_F is an extension by $\mathbb{C}^\times$ of $Gal(\mathbb{C}/\mathbb{R})$ defined by explicit simple formulas (see for example §1.2 of Renard's article in [Book1]), and a local Langlands parameter for G over $\mathbb{R}$ is then a $\hat{G}$-conjugacy class of homomorphisms

$$\phi : W_{\mathbb{R}} \to {}^L G$$

with semisimple image, such that the composite

$$W_{\mathbb{R}} \to {}^L G \to W_{\mathbb{R}} \tag{I.2.1}$$

is the identity.

An admissible representation of $G(F)$ is actually an admissible Harish-Chandra module, which is a complex vector space on which the Lie algebra $Lie(G)$ of G and a maximal compact subgroup of G act compatibly. At least when G is quasi-split over F, Langlands proved a long time ago that local Langlands parameters are in one-to-one correspondence with L-packets of irreducible admissible representations of $G(F)$, where the L-packets were finite collections of representations whose structure was determined by Knapp and Zuckerman. When G is not quasi-split one has to restrict attention to parameters that are *relevant* to G in a precisely defined sense, and then the parametrization remains valid. Missing from this claim is an intrinsic definition of the L-packet, which turns out to be a bit more complicated than one might at first expect. A good way to simplify the parametrization, suggested by Vogan, is to view ϕ as a parameter for all the inner forms of G simultaneously. But for our purposes a rough understanding will suffice.

Now suppose F is non-archimedean, with residue field k of cardinality q. In this case we do need Weil-Deligne parameters. The simplest formulation, and in some ways the most convenient for purposes of comparison, is to use the *local Langlands group* $\mathcal{L}_F = W_F \times SU(2)$, where $SU(2)$ is the familiar Lie group and W_F is the usual Weil group of F, which fits in the following Cartesian diagram:

$$\begin{array}{ccc} W_F & \longrightarrow & Frob^{\mathbb{Z}} \\ \downarrow & & \downarrow \\ \Gamma_F & \longrightarrow & Gal(\bar{k}/k) \end{array}$$

Here $\Gamma_F = Gal(\bar{F}/F)$ and $Frob$ is the standard generator $x \mapsto x^q$. A local Langlands parameter for G is then a $\hat{G}$-conjugacy class of continuous relevant L-homomorphisms

$$\phi : \mathcal{L}_F \to {}^L G \tag{I.2.2}$$

where a homomorphism is called an L-homomorphism if the image in $\hat{G}$ of any element is semisimple and if the obvious analogue of (2.1) is the identity. The word *relevant* is only relevant to us if G is not quasi-split, a situation we will do our best to avoid. When $G = GL(n)$, the *local Langlands correspondence* [HT,He] is a bijection between local Langlands parameters and irreducible admissible representations of $G(F)$. Parameters ϕ that are not trivial on $SU(2)$ arise from interesting issues having to do with the failure of induced representations to be irreducible. For example, when $G = GL(2)$, ϕ is trivial on $SU(2)$ unless the corresponding representation π of $GL(2,F)$ is a Steinberg (special) representation.

When F is non-archimedean but G is not an inner form of $GL(n)$ a given parameter ϕ as in (I.2.2) is conjectured to correspond to a finite packet Π_ϕ of irreducible admissible representations of $G(F)$. This conjecture is now largely proved for classical groups, thanks primarily to the work of Arthur and Moeglin. Again this claim only makes sense when we know how to give a precise meaning to the word "packet". Conjecturally, the elements of Π_ϕ are themselves indexed by irreducible representations of the finite group $\mathfrak{S}_\phi$ of connected components of the centralizer $S_\phi \subset {}^LG$ of the image of ϕ in ${}^LG$. In all cases to be considered here, $\mathfrak{S}_\phi$ is a product of finitely many copies of $\mathbb{Z}/2\mathbb{Z}$. This indexing does not provide an intrinsic characterization of Π_ϕ as a collection of representations of the locally compact group $G(F)$. We return to this issue at the end of the next section.

Assuming the Langlands parametrization, it is easy to define functorial transfer. Any L-homomorphism $\iota : {}^LG' \to {}^LG$ gives rise to a map

$$\{ \text{ parameters for } G' \} \to \{ \text{ parameters for } G \}$$

by the formula

$$\phi' \mapsto \iota_*(\phi') := \iota \circ \phi' \tag{I.2.3}$$

As mentioned above, this gives no intrinsic information about the map on actual representations. Moreover, without the Langlands parametrization one has in general no way of guessing how to associate representations of G to representations of G'. But in the simplest cases of *principal series* representations, this can be worked out explicitly in terms of local class field theory, and the exercise is worth explaining in detail.

I.2.4 Functoriality for principal series representations; the role of Hecke algebras.

Assume G and G' are quasi-split, and let $B = LN$, $B' = L'N'$ be Borel subgroups of G and G' respectively, with L, L' maximal tori and N, N' maximal unipotent subgroups. Let $\chi' : L' \to \mathbb{C}^\times$ be a continuous character. The dual groups $\hat{L}$ and $\hat{L}'$ are maximal tori of $\hat{G}$ and $\hat{G}'$, respectively. Thus the L-homomorphism ι determines a map ${}^LL' \to {}^LL$ and a dual map $\iota^* : L \to L'$ of maximal tori, and we let $\chi = \chi' \circ \iota^*$. The characters χ' and χ define (normalized) induced representations $I(\chi')$ and $I(\chi)$ of G', G respectively. For generic χ', $I(\chi')$ is irreducible, and we want $I(\chi)$ to be an irreducible representation that is the functorial transfer of $I(\chi')$.[2] By class field theory χ' and χ can be interpreted in terms of Weil group characters

$$\phi_{\chi'} : W_F \to {}^LL'; \phi_\chi : W_F \to {}^LL.$$

Composing $\phi_{\chi'}$ with the inclusion ${}^LL' \hookrightarrow {}^LG'$, we obtain a local parameter $\phi_{\chi',G'}$ for G'; likewise ϕ_χ defines a local parameter $\phi_{\chi,G}$ for G, and it is clear that

$$\phi_{\chi,G} = \iota \circ \phi_{\chi',G'}.$$

[2] This is only strictly true if the transferred packet is a singleton; if not the irreducible constituents of $I(\chi)$ will in any case belong to the packet.

The picture is consistent if $\phi_{\chi',G'}$ and $\phi_{\chi,G}$ are taken to be the parameters for $I(\chi')$ and $I(\chi)$, which is valid generically. (On the exceptional set one needs parameters not trivial on the $SU(2)$ factor of $\mathcal{L}_F$.)

Now assume the character χ' is *unramified*, in other words is trivial on the maximal compact subgroup $L'_0 \subset L'$. Then $I(\chi)$ and $I(\chi')$ are *unramified principal series* representations. In particular, letting $K' \subset G'$, $K \subset G$ be *hyperspecial maximal compact subgroups*[3], we have

$$\dim I(\chi)^K = \dim I(\chi')^{K'} = 1.$$

Generically, such an $I(\chi')$ is irreducible and determined by a character of the *unramified Hecke algebra* $\mathcal{H}(G',K')$, which the Satake transform identifies with the polynomial ring $\mathbb{C}[X_{*,F}(L')]^{W'}$, where W' is the Weyl group of G' relative to L' and $X_{*,F}$ is the group of F-rational cocharacters. The map ι^* induces by duality a homomorphism of algebras in the obvious notation

$$\mathcal{H}(G,K)\xrightarrow{\sim}\mathbb{C}[X_{*,F}(L)]^{W} \to \mathbb{C}[X_{*,F}(L')]^{W'}\xrightarrow{\sim}\mathcal{H}(G',K') \tag{I.2.4.1}$$

and thus a map

$$\text{characters of } \mathcal{H}(G',K') \longrightarrow \text{characters of } \mathcal{H}(G,K) \tag{I.2.4.2}$$

or (at least generically) a map from unramified principal series of G' to unramified principal series of G. This gives an explicit construction of functorial transfer for unramified principal series.

I.2.4.3 Integral Hecke algebras. For future reference, we recall that $\mathcal{H}(G,K)$ is the convolution algebra of compactly supported functions on the double coset space $K\backslash G/K$ with identity element 1_K; here it is understood that G is given Haar measure for which K has volume 1. The $\mathbb{Z}$-valued functions in this algebra form a finitely generated $\mathbb{Z}$-subalgebra denoted $\mathcal{H}_{\mathbb{Z}}(G,K)$.

I.3. Global Langlands parameters

Now let F be a number field. The optimal formulation of the Langlands parametrization in this setting is completely parallel to that sketched above for local fields. A Langlands parameter for G is to be a $\hat{G}$-conjugacy class of continuous L-homomorphisms

$$\phi : \mathcal{L}_F \to {}^L G \tag{I.3.1}$$

where $\mathcal{L}_F$ is the *global Langlands group.* The members of the packet Π_ϕ corresponding to ϕ are automorphic representations; in particular they are admissible irreducible representations of the adèle group $G(\mathbf{A}_F)$ which can be represented as restricted tensor products $\pi = \otimes'_v \pi_v$, with each π_v an irreducible admissible representation of $G(F_v)$. For each v there is to be a map

$$\iota_v : \mathcal{L}_{F_v} \to \mathcal{L}_F$$

[3] If G and G' are classical groups defined by bilinear or hermitian forms admitting self-dual lattices, then K and K' are the stabilizers of the corresponding lattices. The general definition of hyperspecial maximal compact is given in [BT] and is unnecessarily technical for this survey.

where for archimedean v, the local Langlands group $\mathcal{L}_{F_v}$ is just the local Weil group, and each π_v is to be an element of the local packet attached to $\phi \circ \iota_v$. By varying the π_v in the local packets, this recipe gives a (possibly infinite) collection of candidates for members of the global packet Π_ϕ.

Define the space of square integrable automorphic forms on G:

$$\mathcal{A}^2(G) = L^2(G(F)\backslash G(\mathbf{A}_F)/Z^0),$$

where Z^0 is the identity component of the product over archimedean places v of $Z_G(F_v)$.[4] This unitary representation of $G(\mathbf{A}_F)$ decomposes as the sum $\mathcal{A}^2(G) = \mathcal{A}_{cont}(G) \oplus \mathcal{A}_{disc}(G)$ of a continuous part and a discrete part, the latter being the Hilbert direct sum of all irreducible subrepresentations of $\mathcal{A}^2(G)$. For the moment, we only consider *discrete automorphic representations*, i.e. those that occur as direct summands of $\mathcal{A}_{disc}(G)$. These are (conjecturally) characterized as those for which the centralizer of ϕ in ${}^LG$ is finite modulo the center of ${}^LG$. Each candidate for membership in Π_ϕ is then a representation of $G(\mathbf{A}_F)$ that occurs as a direct summand in $\mathcal{A}^2(G)$[5] with a certain *multiplicity* (possibly 0) predicted by the Arthur conjectures. (Actually, for Arthur's general multiplicity conjectures one needs not Langlands parameters but rather Arthur parameters, which are (equivalence classes of) L-homomorphisms

$$\phi : \mathcal{L}_F \times SU(2) \to {}^LG \quad \text{(I.3.2)}$$

but the additional complication is beyond the scope of this survey. The simpler formulation will be adequate for our purposes.)

In this setting, functorial transfer of packets is defined exactly as in the local case, by the formula (I.2.3). Compatibility of local and global transfer is automatic, and the multiplicities are governed by Arthur's explicit conjecture. The formulation is optimal, as I wrote above, but has the obvious flaw that the global Langlands group is not an object one knows how to construct but rather the consequence of Langlands' Tannakian formalism [L79], which in turn requires knowledge of functoriality that is by no means on the horizon. For one class of packets one can make precise conjectures, however. In Langlands' original formulation, the hypothetical pro-algebraic group $\mathcal{L}_F$ is expected to map onto the Galois group Γ_F, and so homomorphisms

$$\phi : \Gamma_F \to {}^LG \quad \text{(I.3.3)}$$

compatible with the natural map of the latter to Γ_F are legitimate global Langlands parameters. In this way an n-dimensional Artin representation $\rho : \Gamma_F \to GL(n, \mathbb{C})$ defines a Langlands parameter for $GL(n, \mathbf{A}_F)$, and the Langlands parametrization includes the Artin conjecture as a special case.

In Kottwitz' reformulation, the pro-algebraic $\mathcal{L}_F$ is replaced by a locally compact $\mathcal{L}_F$ that maps to the Weil group W_F rather than the Galois group. This is actually

[4] This normalization of automorphic representations is most convenient for applications of the trace formula but is highly awkward for arithmetic. The arithmetic statements below are not literally true unless an adjustment is made to account for this normalization. We spare the reader the details.

[5] With the appropriate adjustments at the archimedean places.

more convenient in many situations, in particular for local-global comparisons, but it makes no difference for our purposes, and we will not specify which version of the Langlands group we consider.

It's important to mention at this stage, though, that very few Artin Galois representations can be linked to automorphic representations by current methods, and this situation is unlikely to change in the near future. What we can study are the compatible families of ℓ-adic representations arising in the cohomology of Shimura varieties. These are conjecturally linked to the representations of another hypothetical pro-algebraic group, the *motivic Galois group*, the group conjecturally associated to the (conjectural) Tannakian category of motives over F, expected to occur as the middle term in a sequence of surjections

$$\mathcal{L}_F \ \rightarrow \ \Gamma_F^{mot} \ \rightarrow \ \Gamma_F.$$

Since the Tate conjectures imply that the tensor properties of ℓ-adic representations attached to motives faithfully reflect the properties of representations of Γ_F^{mot}, it should in principle be possible to use such ℓ-adic representations as substitutes for global Langlands parameters. However, there is a discrepancy, corresponding to a shift by a half-integral power of the norm character, between Langlands' analytic normalization of the correspondence and the natural motivic normalization; this phenomenon has been analyzed by Buzzard and Gee [BG], who have discovered that Galois representations do not adapt as simply as (some) expected to the framework of the Langlands correspondence.

As a practical substitute for the hypothetical global Langlands parameters, Arthur has proposed to parametrize general automorphic representations of $GL(n)$ by formal sums of the form

$$\psi = \boxplus_{i=1}^{r} (\phi_i, b_i) \tag{I.3.4}$$

where $n = \sum_i n_i$ is an unordered partition, each n_i factors as $n_i = a_i \cdot b_i$ and each ϕ_i is a symbol associated to a cuspidal automorphic representation π_i of $GL(a_i)$.[6] This is consistent with the formalism outlined by Langlands in [L79]. Parameters for one of the classical groups $G' = H_n, J_n, U_n$ considered in §I.1 are then defined to be parameters for $GL(n)$ that satisfy the invariance properties that would be satisfied by a hypothetical Langlands parameter ϕ whose image lies in the image of ${}^LG'$ in ${}^LG_n$. Thus a parameter for U_n is defined to be a parameter as above for $R_{F'/F}(GL(n)_F)$ that is invariant under the outer automorphism defined by the Galois element c,[7] while a parameter for J_n is a parameter (I.3.4) for $GL(2n+1)$ that is self-dual in an appropriate sense; for example, if $r = 1$ and $b_1 = 1$, then the condition is that the representation π_1 be self-dual. The definition of a parameter for H_n is more subtle, to take into account the distinction between symplectic and orthogonal duality in even dimension, but a serviceable definition can be derived from properties of poles of L-functions. One can also express in terms of (I.3.4) the centralizer condition that should imply that a parameter for G' correspond to discrete automorphic representations; we then say that ϕ is *discrete for* G'.

[6] This parametrization implicitly depends on the Moeglin-Waldspurger classification of discrete automorphic representations of $GL(n_i)$ by pairs (ϕ_i, b_i) as above.

[7] There is an additional complication here, corresponding to the fact that there are two ways U_n can occur as a twisted endoscopic group for $R_{F'/F}(GL(n)_F)$. Fortunately, only one of these is relevant for the representations we consider.

In their work on the local Langlands correspondence for classical groups, Arthur and Moeglin define parameters in a similar way. Since the local Langlands correspondence is known for $GL(n)$, however, their parameters can be defined as homomorphisms to the L-group of the appropriate $GL(n)$ that lie in the L group of the classical group. We return to this question in (II.4.3).

Since discrete automorphic representations of the groups H_n, J_n, U_n of I.1 often transfer to representations of $GL(n)$ that are not discrete, one cannot avoid considering partitions of n as in (I.3.4). For example, consider an Artin parameter ϕ as in (I.3.3) with values in ${}^L H_n$. If its centralizer in ${}^L H_n$ is finite, then ϕ is a discrete parameter for H_n, hence is expected to determine a discrete L-packet Π_ϕ for $H_n(\mathbf{A}_F)$. But it is quite possible that the composition of ϕ with the inclusion of ${}^L H_n$ in ${}^L GL(2n)$ defines a reducible $2n$-dimensional representation; then the transfer of Π_ϕ will not be discrete on $GL(2n, \mathbf{A}_F)$. Other examples will be given in §II.5.

I.3.5. Tensor product functoriality. The Langlands functoriality conjectures, applied to $GL(n)$, include the following prediction.

Conjecture. *Let*

$$\tau : GL(n_1) \times GL(n_2) \times \ldots GL(n_r) \longrightarrow GL(N)$$

be an irreducible algebraic representation. Let F be a number field and let $\pi_1, \ldots, \pi_r$ be cuspidal automorphic representations of $GL(n_i, F)$, $i = 1, \ldots, r$. Then there is a (necessarily unique) automorphic representation (functorial transfer) $\tau(\pi_1 \boxtimes \pi_2 \boxtimes \cdots \boxtimes \pi_r)$ of $GL(N, F)$ such that, at almost all places v,

$$\mathcal{L}(\tau(\pi_1 \boxtimes \pi_2 \boxtimes \cdots \boxtimes \pi_r)_v) = \tau \circ (\mathcal{L}(\pi_{1,v}) \otimes \cdots \otimes \mathcal{L}(\pi_{r,v})),$$

where $\mathcal{L}(\pi_{i,v})$ is the n_i-dimensional representation of the local Langlands group $\mathcal{L}_{F_v}$ given by the local Langlands correspondence.

Whether or not the transfer is cuspidal depends both on τ and on the original π_i, and the answer does not admit a simple description. However, when $r = 1$, one expects that its transfer will be cuspidal provided $\pi = \pi_1$ is sufficiently general; then the standard L-function $L(s, \tau(\bullet))$ is *entire* and satisfies the functional equation for $GL(n)$. In the case of automorphic representations of $GL(2)$ attached to elliptic curves, " sufficiently general" excludes only elliptic curves with complex multiplication, cf. §III.3.2.

The case $r = 2$, with $N = n_1 \cdot n_2$ and τ the standard tensor product representation, may be considered the main open question in automorphic forms. The following list exhausts all the cases I know, excluding the trivial cases, where some of the $n_i = 1$):

(1) $r = 1$, $n(= n_1) = 2$, $N = 3$ (symmetric square): [GJ]
(2) $r = 1$, $n(= n_1) = 2$, $N = 4, 5$ (symmetric cube and fourth power) [KS, Ki]
(3) $r = 1$, $n(= n_1) = 4$, $N = 6$ (exterior square) [Ki]
(4) $r = 2$, $n_1 = n_2 = 2$, τ the tensor product: [Ra].
(5) $r = 2$, $n_1 = 2$, $n_2 = 3$, τ the tensor product: [KS]

These results are *unconditional* and apply to all automorphic representations over all number fields. Arithmetic arguments have been developed to prove results for higher symmetric powers for $GL(2)$ that are much weaker than the above conjecture, but suffice to prove the Sato-Tate Conjecture in many cases. These will be discussed in §III.3.

Having defined parameters for a classical group G' by reference to the formal parameters (I.3.4), Arthur's task is then to show that any parameter for $GL(n)$ that (formally) "factors through" the L-group of G' does in fact define an automorphic packet for the latter, at least when the parameter is discrete for G', and that all discrete automorphic representations of G' arise in this way. The proof, as of this writing still in progress (to appear in [A10], but see [A05, §30]) is naturally based on the stable trace formula, to which we now turn.

Part II. Stabilization of the trace formula

1. Introduction to the simple trace formula

Let G be a connected reductive group over the number field F. Let $C_c^\infty(G(\mathbf{A}))$ denote the space of *test functions*: compactly supported functions on $G(\mathbf{A}_F)$ that are C^∞ in the archimedean variables[8] and locally constant in the non-archimedean variables. These act as integral operators on the space $\mathcal{A}^2(G)$: given $v \in \mathcal{A}^2(G)$, $f \in C_c^\infty(G(\mathbf{A}))$, define

$$R(f)v(h) = \int_{G(\mathbf{A}_F)} f(g)v(hg)dg, \tag{II.1.1}$$

where dg is a fixed Haar measure on G; we will always take *Tamagawa measure*. The (Arthur-Selberg) *trace formula* is an expression for the trace of the representation R_{disc} induced by (II.1.1) on the subspace $\mathcal{A}_{disc} \subset \mathcal{A}^2(G)$, in terms intrinsic to the geometry of G. By *trace* we mean the distribution on $G(\mathbf{A})$ that, to a function $f \in C_c^\infty(G(\mathbf{A}))$ assigns the number

$$T_{disc}(f) := tr\ R_{disc}(f),$$

well-defined because of the (by no means trivial) fact that $R_{disc}(f)$ is a trace class operator. A related non-trivial fact is that there is a Hilbert space decomposition

$$\mathcal{A}_{disc} = \oplus_\pi m_\pi \pi \tag{II.1.2}$$

where π runs over all unitary irreducible representations of $G(\mathbf{A}_F)$ and the multiplicities m_π are finite non-negative integers, positive only for a countable set of π.

Suppose for the moment that G is anisotropic; then $\mathcal{A}_{disc} = \mathcal{A}^2(G)$ and the trace formula admits a simple expression. Suppose $\gamma \in G(\mathbf{A})$ and f is a test function. Let $H_\gamma \subset G(\mathbf{A})$ denote the centralizer of γ. Suppose for the moment that H_γ has a Haar measure dg_γ, and let $d\dot{g} = dg/dg_\gamma$ denote the quotient measure on $H_\gamma\backslash G(\mathbf{A})$. We define the orbital integral

$$O_\gamma(f) = \int_{H_\gamma\backslash G(\mathbf{A})} f(x^{-1}\gamma x)d\dot{x}. \tag{II.1.3}$$

[8] One also needs to assume that they transform with respect to a finite-dimensional representation under the right and left actions of a fixed maximal compact subgroup of the archimedean part of $G(\mathbf{A})$.

Convergence of the integral is automatic for anisotropic G if $\gamma \in G(F)$. In that case, H_γ is the group of $\mathbf{A}_F$-points of the centralizer G_γ of γ, an F-rational reductive algebraic subgroup of G, and we let dg_γ denote Tamagawa measure again. The map $f \mapsto O_\gamma(f)$ defines a ($G(\mathbf{A})$-)invariant distribution on $C_c^\infty(G(\mathbf{A}))$. On the other hand, dg_γ defines a volume (the Tamagawa number)

$$\tau_\gamma = \int_{G_\gamma(F)\backslash G_\gamma(\mathbf{A}_F)} dg_\gamma.$$

Theorem II.1.4 (Selberg trace formula, anisotropic case). *There is an identity of distributions:*

$$T_{disc}(f) = \sum_\gamma \tau_\gamma O_\gamma(f)$$

where γ runs over $G(F)$-conjugacy classes in $G(F)$ and, for any fixed choice of f, the sum of orbital integrals is finite. In other words

$$\sum_\gamma \tau_\gamma O_\gamma(f) = \sum_\pi m_\pi tr\ \pi(f) \tag{II.1.5}$$

where the m_π are as in (II.1.2).

The left-hand side of (II.1.5) is called the *geometric side* of the trace formula, the right-hand side is the *spectral side*; the distribution $f \mapsto tr\ \pi(f)$ is the *distribution character* of the representation π, one of the fundamental concepts in Harish-Chandra's approach to representation theory of Lie groups. The formula may well appear mysterious at first sight, but once the analytic issues are resolved, and they are not so problematic for anisotropic groups, the proof is quite intuitive; those unfamiliar with the formula are encouraged to read Bump's article [Bump].

Matters are quite different when G is not anisotropic, where the general formula is due to Arthur. While there is again an equality of geometric and spectral sides, and the terms corresponding to *elliptic* conjugacy classes γ and *discrete* π resemble the corresponding terms in (II.1.4), there are also more complicated terms coming from the remaining conjugacy classes in $G(F)$ on the left and from the continuous spectrum on the right. Arthur has developed a *simple trace formula* in which these problematic terms all vanish, and for the purposes of the present exposition they can be ignored, but for the most general results, including the crucial Arthur-Clozel base change theorem and Arthur's recent work on functorial transfer from classical groups, use of the complete Arthur trace formula cannot be avoided.

By far the best survey of the trace formula is Arthur's introduction [A05], which has been frequently consulted in the preparation of the present article.

II.1.6. The twisted trace formula.

We will see below that the most general results in the direction of functoriality have been proved by comparing the geometric sides of trace formulas for the two groups involved. In some of the most important cases, the group denoted G in the comparison, where the L-homomorphism is of the form $\iota : {}^LG' \to {}^LG$, is replaced by the disconnected semidirect product of G with a finite cyclic group Θ of automorphisms. For example, L/F is a cyclic extension of prime degree with Galois group Θ, $G = R_{L/F}G'_L$, and ι is the natural (diagonal) imbedding, the corresponding functoriality is called base change, and has been established in a number of situations by comparing the trace formula for G' with the trace formula for $G \rtimes \Theta$. The trace formula for such a disconnected group is called the *twisted trace formula*; its study when $G = GL(2)$ was pioneered by Saito and Shintani before being developed systematically by Langlands (for $GL(2)$) and Arthur-Clozel (for general $GL(n)$). Abstract treatments of issues connected with the twisted trace formula include the books of Kottwitz-Shelstad [KSh] and Labesse [Lab99] as well as unpublished notes of Clozel, Labesse, and Langlands dating from 1983-84 (the "Morning Seminar on the Trace Formula" at the IAS, currently being rewritten in detail by Clozel and Labesse).

For our purposes here, it suffices to mention that the twisted trace formula, like the formula (II.1.5), is an equality between a geometric side and a spectral side. The spectral side is roughly the sum over discrete[9] automorphic representations π that are invariant under the action of Θ on $G(\mathbf{A})$, whereas the geometric side is a sum over conjugacy classes γ as before of *twisted orbital integrals*, which incorporate the action of a chosen generator of Θ. We will not need the explicit form of this formula (but see [A05,§30]).

II.1.7. Cohomological representations.

Assume for the moment that G is anisotropic. Then Z^0 is trivial and $\mathcal{A}_{disc} = \mathcal{A}^2(G)$ is the space of square-integrable functions on the *compact* space $\tilde{S}(G) = G(F)\backslash G(\mathbf{A}_F)$, a projective limit of (disconnected) compact manifolds. Let

$$K_\infty \subset G_\infty := \prod_{v \mid \infty} G(F_v)$$

be a maximal connected compact subgroup, so that

$$S(G) = \tilde{S}(G)/K_\infty \qquad \text{(II.1.7.1)}$$

is a projective limit, over open compact subgroups of the group $G(\mathbf{A}^f)$ of finite adèles of G, of locally symmetric spaces, each isomorphic to a finite union of manifolds of the form $\Delta\backslash G_\infty/K_\infty$ where Δ varies over arithmetic subgroups of G_∞. When G_∞/K_∞ is of hermitian type, $S(G)$ is a *Shimura variety* which in the anisotropic case is a projective limit of complex projective varieties with a canonical model over a certain number field, the *reflex field*.

For appropriate choices of test functions f, $T_{disc}(f)$ can be identified with the trace of a correspondence, acting on the cohomology of the pro-manifold $S(G)$ with coefficients in a local system. Let W be a finite-dimensional algebraic representation of G and for any open compact subgroup $K_f \subset G(\mathbf{A}^f)$, let

$$\tilde{W}_{K_f} = G(F)\backslash G(\mathbf{A}_F) \times W/K_f \cdot K_\infty \qquad \text{(II.1.7.2)}$$

be the corresponding local system over ${}_{K_f}S(G) := S(G)/K_f$. Assume $f = f_\infty \otimes f_f$, corresponding to the factorization $G(\mathbf{A}_F) \xrightarrow{\sim} G_\infty \times G(\mathbf{A}^f)$, and assume $f_f(kgk') = f_f(g)$ for all $k, k' \in K_f$. Then f_f defines a generalized Hecke correspondence $H(f_f)$ on ${}_{K_f}S(G)$ by the standard integral formulas (using a factorization of our chosen Tamagawa measure on $G(\mathbf{A}_F)$ as product of archimedean and finite parts) that

[9] More generally, Θ-discrete, in a sense that will not be defined here.

acts naturally on the cohomology of any $\tilde{W}_{K_f}$. Moreover, there is a class of test functions $f_\infty = EP_W$, called *Euler-Poincaré functions*, with the property that

$$T_{disc}(EP_W \otimes f_f) = \sum_{i \geq 0} (-1)^i Tr(H(f_f) \mid H^i({}_{K_f}S(G), \tilde{W}_{K_f})). \quad \text{(II.1.7.3)}$$

This is a consequence of Matsushima's formula: defining the multiplicities $m(\pi)$ as in (II.1.2), and factoring $\pi = \pi_\infty \otimes \pi_f$,

$$H^i({}_{K_f}S(G), \tilde{W}_{K_f}) \xrightarrow{\sim} \oplus_\pi \, m(\pi) H^i(Lie(G), K_\infty; \pi_\infty \otimes W) \otimes \pi_f^{K_f}, \quad \text{(II.1.7.4)}$$

where $H^i(Lie(G), K_\infty, \bullet)$ is relative Lie algebra cohomology [BW]. The representations π_∞ for which $H^i(Lie(G), K_\infty; \pi_\infty \otimes W) \neq 0$ are called *cohomological* (for W).

When G is not anisotropic there are technical complications. Care needs to be taken with the center of G; in particular, not every W defines a local system. A more substantial complication is that (II.1.7.3) calculates the trace on L^2-cohomology rather than cohomology. When $S(G)$ is a Shimura variety (cf. §III.1), Zucker's conjecture (proved over twenty years ago by several methods) affirms this is isomorphic to the cohomology of the Baily-Borel-Satake compactification ${}_{K_f}S(G)^*$ of ${}_{K_f}S(G)$, with coefficients in the middle perverse extension $j_{!*}(\tilde{W}_{K_f})$, and this is what replaces the left-hand side of (II.1.7.4) in this case.

On the other hand, the Lefschetz formula expresses the trace of the correspondence $H(f_f)$ on the cohomology of $\tilde{W}_{K_f}$ as a sum of orbital integrals, and in the anisotropic case one finds that this is equivalent to the identity (II.1.5). In general, Arthur has shown in [A89] that (II.1.5) expresses the Lefschetz number of the correspondence $H(f_f)$ on L^2-cohomology; the corresponding result for middle intersection cohomology was proved by Goresky, Kottwitz, and MacPherson by a topological argument in [GKM]. Using a new interpretation of the intersection complex when $S(G)$ is a Shimura variety, S. Morel [Mo] was able to incorporate the action of Frobenius at a chosen finite place of the reflex field.

II.2. Langlands' program for endoscopy

The distributions that appear on the two sides of (II.1.5) can be considered the basic objects in the harmonic analysis on the adèlic group $G(\mathbf{A})$. The distribution character $tr\ \pi(f)$ is the analogue of the character of an irreducible representation of a finite group, whereas the orbital integral is the analogue of averaging over a conjugacy class. In particular, both expressions in (II.1.5) are sums of *invariant* distributions, in other words distributions that are invariant under conjugation by the group.

The geometric side is so-called because each term is an integral over the orbit of the action of the group $G(\mathbf{A})$ on itself by conjugation, itself the restricted direct product of the local orbits of the groups $G(F_v)$. Such an orbit is not always "geometric" in the sense of algebraic geometry, however. The set of F_v-rational points of the $G(\bar{F}_v)$-orbit of a semisimple element $\gamma \in G(F_v)$ is a single orbit when $G = GL(n)$, or an inner twist of $GL(n)$. Concretely, a semisimple conjugacy class in $GL(n, F_v)$ is characterized by its characteristic polynomial, and if two semisimple elements of $GL(n, F_v)$ are conjugate over $\bar{F}_v$, then they have the same characteristic polynomial, hence are already conjugate over F_v. In contrast, for general groups the intersection of the $G(\bar{F}_v)$-orbit of γ with $G(F_v)$ can be a finite union of distinct $G(F_v)$-conjugacy classes.

A distribution on test functions on $G(F_v)$ (or $G(\mathbf{A})$) is called *stable* if it is invariant under *stable conjugacy*, which in the cases of primary interest to us means conjugacy by elements of $G(\bar{F}_v)$. An example is the *stable orbital integral*

$$SO_\gamma(f) = \sum_{\gamma' \overset{st}{\sim} \gamma} O_{\gamma'}(f), \tag{II.2.1}$$

where $\gamma' \overset{st}{\sim} \gamma$ if γ' is stably conjugate to γ, the sum taken over stable conjugacy modulo conjugacy. A crucial complicating feature of the trace formula is that, except for $GL(n)$ and its inner forms, the distribution T_{disc} is **not** a sum of stable distributions. This is in principle an obstacle to using the trace formula to prove functoriality, since the first step in the comparison of the geometric sides of trace formulas for groups G, G' is invariably the comparison of (semisimple) conjugacy classes of G and G' in terms of their geometric invariants.

The failure of the trace formula to be stable was first encountered for the group $SL(2)$ [LL]; the authors discovered that the geometric side of the trace formula for $SL(2)$ could be written as a sum of a stable distribution (the sum of *stable orbital integrals* for $SL(2)$), which we will denote ST_{disc}, and an error term expressible in terms of Galois cohomology. After a lengthy manipulation, the error could be related to the trace formulas for a collection of smaller auxiliary groups, now called *endoscopic groups*. In his book [L83] Langlands proposed an expected generalization valid for all G. Nearly thirty years later, this generalization has now been completely established. The names of the protagonists will be revealed in the next few sections.

In the cases of interest, the endoscopic groups attached to G are groups H, always assumed *quasi-split*, with the property that ${}^LH$ is the centralizer of an element of order 2 $s \in \hat{G}$. More precisely, one considers *elliptic endoscopic data* (H, s, ξ) where H and s are as indicated and $\xi : {}^LH \to {}^LG$ is an L-homomorphism.[10] The set of such (H, s, ξ) is taken up to an appropriate equivalence relation and depends only on the class of G up to inner automorphism. For the cases introduced in §1, there are only finitely many equivalence classes, easily described in terms of linear algebra:

(II.2.2.1) For $G = GL(n)$ or an inner twist, the only (elliptic) endoscopic group is G itself.

(II.2.2.2) For $G = H_n$, the elliptic endoscopic groups are of the form $H_k \times H_j = SO(2k+1) \times SO(2j+1)$ where $2k + 2j = 2n$ and $k \leq j$. When $k = 0$, $H_k \times H_j$ is just H_n itself; this *principal endoscopic group* corresponds to the "stable part" of the trace formula for H_n.

(II.2.2.3) For $G = U_n$, the elliptic endoscopic groups are of the form $U_k^* \times U_j^*$ with $k + j = n$, $k \leq j$. The superscript * indicates that we are taking *quasi-split* unitary groups. The case $k = 0$ yields the quasi-split inner form of G itself, the principal endoscopic group.

(II.2.2.4) For $G = J_n$, the elliptic endoscopic groups are of the form $SO(2k)^* \times Sp(2j)$, with $2k + 2j = 2n$. As above * denotes a quasi-split form, and $k = 0$ is the principal endoscopic group.

[10] In the most general case the source of the map ξ is not ${}^LH$ itself but a central extension of ${}^LH$, but this suffices for a first approximation.

In (3), the identification of an elliptic endoscopic group with a classical group depends crucially on the choice of L-homomorphism ξ. We return to this point when we discuss the case of U_n below.

Fix G anisotropic for the moment, and let $\mathfrak{E}$ denote the set of equivalence class of elliptic endoscopic data for G. The stable trace formula for G should take the form

$$T_{disc}(f) = \sum_{(H,s,\xi)\in\mathfrak{E}} i(G,H) ST^H_{disc}(f^H). \tag{II.2.3}$$

The constants $i(G,H)$ are expressions involving Tamagawa numbers and group-theoretic data; thanks primarily to Kottwitz [K84, K86, K88] they are easy to calculate. The expression ST^H_{disc} designates a stable distribution on the endoscopic group H. It should be thought of to a first approximation as the sum of the stable orbital integrals of the test function f^H over elliptic conjugacy classes in $H(\mathbf{A})$, but since H is quasi-split this is in general insufficient, and the correct definition of the term ST^H_{disc} is only obtained by Arthur at the end of a lengthy and difficult induction process. We describe this briefly in §II.4 below. The reader will meanwhile have noticed a more immediate problem with (II.2.3), namely that the right-hand side depends on a collection of test functions $\{f^H\}$ that we have not pretended to define.

The real challenge of endoscopy is to define the *transfer maps* $f \mapsto f^H$. These are local in the sense that, if $f = \otimes' f_v$ factors over places of v, then f^H can be taken in the form $\otimes' f_v^H$ where $f_v \mapsto f_v^H$ is a local transfer map. These need to satisfy two properties, *a priori* unrelated. The first is the equality (II.2.3). Since the term ST^H_{disc} appearing on the right is a stable distribution, the test function f_v^H only needs to be determined as a functional on stable distributions, and it turns out that it suffices to specify its stable orbital integrals on regular semisimple conjugacy classes in $H(F_v)$. For elliptic elements, this is given by an explicit formula obtained by Langlands and, more generally, by Kottwitz, by analyzing the failure of the geometric side of the trace formula to be stable.

The second property is suggested by the fact that the endoscopic group H is given along with an L-homomorphism

$$\xi : {}^L H \to {}^L G.$$

By the principle of functoriality, this should correspond to a functorial transfer map from (local or automorphic) representations of H to representations of G. Now a representation is determined by its distribution character, and in this way representations can be viewed as dual objects to test functions. Thus the second property is that the map $f_v \mapsto f_v^H$ be dual to the functorial transfer of representations from $H(F_v)$ to $G(F_v)$. Since f_v^H is only determined as a functional on stable distributions, it follows heuristically that the source of functorial transfer should be a collection of representations of $H(F_v)$, the sum of whose distribution characters is a stable distribution on H. This collection is exactly what is meant by a packet, and the proof of existence of finite packets in this sense is inseparable from the stabilization of the trace formula.

The compatibility of these two properties for transfer of packets corresponding to *unramified* local parameters, as described at the end of §I.2, is precisely the *Fundamental Lemma.* We address this issue in detail in §II.3.

II.2.4. Twisted endoscopy.

The most striking examples of functorial transfer have been established by means of a variant of endoscopy, called *twisted endoscopy*, adapted to the stabilization of the twisted trace formula. Given a twisted group $G \rtimes \Theta$ as in II.1.6, its trace formula should also admit a stable version analogous to (II.2.3), where $\mathfrak{E}$ now runs over the set of *twisted endoscopic data* for $G \rtimes \Theta$. These are also triples of the form (H, s, ξ), where H is (fortunately) a connected reductive group over F.

When $G' = GL(n)_F$ and $G = G'_L = GL(n)_L$, as in the base change situation discussed in (II.1.6), G' is the sole (twisted) endoscopic group for $G \rtimes \Theta$, and the analogue of (II.2.3) is the comparison of the twisted trace formula for G with the trace formula for G' developed in [L80] and [AC]. When $G' = U_n$ relative to a quadratic extension F'/F, as in §I.1, $G = G'_{F'}$, and $\Theta = Gal(F'/F)$, then we are again in a base change situation and G' is the principal endoscopic group for $G \rtimes \Theta$, but there are other endoscopic groups as well. Indeed, the endoscopic groups in this case are of the form $U(a) \times U(b)$ with $a+b = n$, and the group G' occurs twice with different data (s, ξ). Arthur uses the terminology *primitive elliptic endoscopic data* rather than "principal endoscopic groups," and we do the same in the following discussion.

Now let $G = GL(2n+1)$ and let Θ be the group of order two generated by the outer automorphism $\theta : g \mapsto {}^t g^{-1}$. The twisted endoscopic groups for $G \rtimes \Theta$ are the split groups $SO(2k+1) \times Sp(2(n-k))$), for $0 \leq k \leq n$, and the datum ξ involves the choice of a quadratic Hecke character η of F. The primitive endoscopic data correspond to the case $k = 0$; the group is the one denoted J_n above. In this case the centralizer $\hat{G}_s$ of the second component $s \in \hat{G}$ of the endoscopic datum is the disconnected group $O(2n+1) \xrightarrow{\sim} SO(2n+1) \times \{\pm 1\}$. The quadratic character η defines a map $W_F \to \{\pm 1\} \subset \hat{G}_s$, and each choice of η defines a distinct L-embedding $\xi = \xi_\eta$, and thus a distinct primitive endoscopic datum. Thus the analogue of (II.2.3) compares the twisted trace formula for G with the stable trace of J_n, or rather the sum over varying η; this comparison is the basis of Arthur's proof of functorial transfer in this case.

If now $G = GL(2n)$, with Θ as above, the twisted endoscopic groups for $G \rtimes \Theta$ are the quasi-split groups $SO(2k, \eta) \times SO(2(n-k)+1)$, for $0 \leq k \leq n$, where η is again a quadratic Hecke character that determines a quasi-split outer form of the even orthogonal group. In this case there are **two** kinds of primitive endoscopic data, corresponding to $k = 0$, the group denoted H_n above, and $k = n$, where one again finds an infinite family parametrized by the additional datum η. The twisted trace formula here is used by Arthur to establish functorial transfer to $GL(n)$ for even as well as odd orthogonal groups. See §30 of [A05], as well as §II.4 below, for more details.

II.3. The Fundamental Lemma and local transfer

In this section F is a *local field*.

Recall the two properties required of the transfer map:

(II.3.1) Equality (II.2.3) on the geometric side (*normalization by orbital integrals*)
(II.3.2) Compatibility with local functoriality (*normalization by characters*)[11]

[11] For unramified parameters, this comes down to compatibility with the relation I.2.4.2.

That it is possible to satisfy either of these requirements separately is already a deep statement in local harmonic analysis. That the two properties are satisfied simultaneously confirms the basic compatibility of Langlands' functoriality with representation theory. The two properties were first confirmed for real groups in a series of papers by Shelstad [Shel79,Shel82], which provided the model for the general local conjectures.

The local equality underlying (II.2.3) is known as the *Langlands-Shelstad transfer conjecture*, now a theorem in almost all situations. Before discussing the transfer of local test functions we need to explain the transfer of local conjugacy classes. The main point is that, although the endoscopic group H is not generally isomorphic to a subgroup of the group G, the two groups do share certain classes of maximal tori. More precisely, one can find pairs of maximal tori $(T^H \subset H, T \subset G)$ such that $\iota : T^H \xrightarrow{\sim} T$ compatibly with the embedding of ${}^L H$ in ${}^L G$ (cf. [A05, p. 199], [H2,§5]). An element $\delta \in T^H$ such that $\iota(\delta) = \gamma$ is called an *image* from G, or a *norm* of γ.

In this correspondence only the *stable conjugacy class* of δ is well-defined relative to γ. The correspondence is most readily illustrated when $G = U_n$ and $H = U^*_{n_1} \times U^*_{n_2}$ with $n_1 + n_2 = n$, as discussed above. For simplicity we assume $G = G^*$ is quasi-split. In this case H can be identified with a subgroup of G; the reader should be aware that the identification depends on the additional datum ξ, but we will not worry about this. Say $G = U(V)$ for a maximally isotropic hermitian space V relative to the quadratic extension F'/F. A typical torus of G is obtained by choosing a collection $\{(F_i, V_i, i \in I), j\}$ where $[F_i : F] = n_i$, $\sum_{i\in I} n_i = n$, V_i is a one-dimensional hermitian space over $F'_i = F' \otimes_F F_i$, and

$$\alpha : \ \oplus_i R_{F'_i/F'} V_i \xrightarrow{\sim} V \tag{II.3.3}$$

is an isomorphism of hermitian spaces over F'. Via α the group $T := \prod_{i\in I} U(V_i)$ is identified with a maximal torus in G – let j_α denote the inclusion of T in G – which is elliptic provided no F_i contains F', which we now assume. An element $t_i \in U(V_i)$ can be viewed as an element of $(F'_i)^\times$, and $t = (t_i) \in T$ is *elliptic regular* provided each t_i generates F'_i over F' and no pair (t_i, t_j) can be identified via field isomorphisms $F'_i \xrightarrow{\sim} F'_j$. Note that the group T depends up to isomorphism only on the fields F_i and not on the choice of hermitian spaces (there are two one-dimensional hermitian spaces over F'_i but the corresponding unitary groups are isomorphic). Then the number of conjugacy classes in the stable conjugacy class of $\gamma := j_\alpha(t)$ is $2^{|I|-1}$, which is the number of possible choices of hermitian spaces V_i, modulo simultaneous scaling of the hermitian forms by the same factor in $F^\times$.

With these identifications, the transfer of conjugacy classes to endoscopic groups is easy to describe. The torus T embeds in the endoscopic group $H = U^*_{n_1} \times U^*_{n_2}$ provided there is a partition $I = I_1 \coprod I_2$ such that $n_j = \sum_{i\in I_j}, j = 1, 2$. The partition determines an obvious factorization $T = T_1 \times T_2$ and a family of embeddings $j_\beta : T \hookrightarrow H$ well-defined up to stable conjugacy. The image $\delta = j_\beta(t)$ is then a norm of γ in H, and we write $\gamma \circ\!\!\rightarrow \delta$, adapting notation introduced by Labesse to emphasize that the norm is only defined as a stable conjugacy class. Stable conjugacy and endoscopic transfer can be expressed in similar elementary terms for other classical groups; see [W01, §10] for details.

Transfer can be defined more generally for non-elliptic semi-simple conjugacy classes, using Harish-Chandra's method of descent. We can now state a precise version of requirement (II.3.1):

II.3.4. Langlands-Shelstad Transfer Conjecture. *Let F be a local field. Let (H, s, ξ) be an elliptic endoscopic triple for G. Let $f \in C_c^\infty(G(F))$. Then there is a function $f^H \in C_c^\infty(H(F))$, uniquely determined as a functional on stably invariant distributions on H, such that, for δ a stable semi-simple conjugacy class in H and γ a regular semi-simple conjugacy class in G, with $\gamma \circ\!\!\!\rightarrow \delta$, we have*

$$SO_\delta(f^H) = \sum_{\gamma' \overset{st}{\sim} \gamma} \Delta(\delta, \gamma') O_{\gamma'}(f). \tag{II.3.5}$$

We write $f \circ\!\!\!\rightarrow f^H$, following Labesse. In the equality (II.3.5) $\gamma' \overset{st}{\sim} \gamma$ denotes stable conjugacy, as above, and the term $\Delta(\delta, \gamma')$ is the notorious *transfer factor* defined generally in [LS87,LS90], that equals 0 unless $\gamma \circ\!\!\!\rightarrow \delta$. The transfer factor is a product of several terms, the most important of which measures the Galois cohomological obstruction that differentiates conjugacy from stable conjugacy. The other terms are more mysterious. Simplifications due to Hales, Waldspurger, and Kottwitz (cf. [K99]) eliminate these terms in favorable situations, and the transfer factor used in the proof of the Fundamental Lemma is quite elementary. But the other terms cannot be avoided in the general situation. Their complexity can be gauged by the fact that the most important step in the article [W08], to which we return below, is the proof of an equality of transfer factors that occupies 150 pages.

We now turn to (II.3.2). Fix the endoscopic datum (H, s, ξ), and let D be a stable distribution on H. Formula (II.3.5) implies that the map

$$f \mapsto D(f^H)$$

defines an invariant distribution on G, which we denote $\xi_*(D)$. Suppose now that

$$\phi : \mathcal{L}_F \to {}^L H$$

is the local Langlands parameter corresponding to a packet Π_ϕ. Let D_ϕ be the stable character attached to the packet.[12] Condition (II.3.2) is then that $\xi_*(D_\phi)$ should have a "natural" expression in terms of the packet $\xi_*(\phi)$, defined as in §I.2. This expression will be a sum of the elements of $\xi_*(\phi)$ with integer coefficients that can be considered the spectral analogues of the transfer factors used in the statement of (II.3.5).

Thanks to the work of Harish-Chandra and Langlands, functoriality for representations of real groups could be expressed in terms of local parameters and Shelstad could check properties (II.3.1) and (II.3.2) simultaneously, defining spectral transfer factors in the process. Over non-archimedean local fields only the stable orbital integrals can be defined unambiguously; parametrization is only available initially for principal series representations, as explained in §I.2.4. Assume G and H are quasi-split, as above, and assume they split over an unramified extension of F. Let $K \subset G$, $K_H \subset H$ be hyperspecial maximal compact subgroups. We consider

[12]This is not a simple notion in general. If ϕ is a *tempered* parameter, then D_ϕ should just be the sum of the distribution characters of the members of the packet with coefficients all equal to 1 for the classical groups introduced in §I.1. If not, the Langlands parameter needs to be replaced by an Arthur parameter, and the arguments that follow need to be modified to take this into account.

functoriality for irreducible *unramified principal series* π. The set of such representations π of H is in one-to-one correspondence with (generic) characters of the Hecke algebra $\mathcal{H}(H, K_K)$, and functoriality is given by the map (I.2.4.2), dual to the map (I.2.4.1) on Hecke algebras. Property (II.3.2) is then guaranteed, in the case of unramified principal series, by the following conjecture of Langlands and Shelstad:

II.3.6. Fundamental Lemma for unramified Hecke algebras. *The restriction of the map $f \multimap f^H$ to $\mathcal{H}(G, K)$ is the map*

$$\xi^* : \mathcal{H}(G, K) \to \mathcal{H}(H, H_K)$$

of (I.2.4.1).

Recall that $\multimap$ is not a real map; the Fundamental Lemma is thus the identity

$$SO_\delta(\xi^*(f)) = \sum_{\gamma' \overset{st}{\sim} \gamma} \Delta(\delta, \gamma') O_{\gamma'}(f) \ \forall f \in \mathcal{H}(G, K) \tag{II.3.7}$$

When $f = 1_K$ is the characteristic function of K, which is the identity element in $\mathcal{H}(G, K)$, (II.3.7) becomes

II.3.8. Fundamental Lemma for unit elements.

$$SO_\delta(1_{K_H}) = \sum_{\gamma' \overset{st}{\sim} \gamma} \Delta(\delta, \gamma') O_{\gamma'}(1_K)$$

This is the statement that has come to be known simply as "the Fundamental Lemma," or more appropriately, "the Fundamental Lemma for endoscopy." It is more fundamental than the others for several reasons.

Theorem II.3.9 (Hales, [Ha95]). *Assume there is an integer N such that (II.3.8) holds for all p-adic fields when $p > N$. Then (II.3.6) holds for all p-adic fields.*

This is proved by a global argument of a sort that has become increasingly familiar in the field, obtaining information about local harmonic analysis by applying special cases of the simple trace formula. By a similar but more elaborate global argument, Waldspurger obtained the following rather surprising result.

Theorem II.3.10 (Waldspurger, [W97]). *Assume the Fundamental Lemma (II.3.8) for all p-adic fields. Then the Langlands-Shelstad transfer conjecture is valid for all p-adic fields.*

Following the proof of special cases by Goresky, Kottwitz, and MacPherson, and work of Laumon, Laumon and Ngô proved (II.3.8) for unitary groups [LN] over local fields of sufficiently large *positive characteristic.* Generalizing and extending the method of [LN], Ngô then proved the following theorem:

Theorem II.3.11 [N06,N2]. *Assume F is a field of positive characteristic p that is sufficiently large relative to the group G. Then the Fundamental Lemma (II.3.8) holds for G and any of its elliptic endoscopic groups H.*

Together with the reductions (II.3.9) and (II.3.10), this practically solves the problem of local transfer, thanks to this result, proved in two ways at roughly the same time:

Theorem II.3.12 [W3,CHLo]. *Let G be a reductive group over $\mathbb{Z}$ and let p be a prime number sufficiently large relative to the rank of G [not too large, in fact]. Let F and F' be two non-archimedean local fields with residue field k of characteristic p. Then the Fundamental Lemma (II.3.8) is valid for G_F if and only if it is valid for $G_{F'}$.*

The articles [W3] and [CHLo] actually prove a version of Theorem II.3.12 for Lie algebras rather than groups, but an argument due to Waldspurger shows this is sufficient. The principles of the two proofs are quite different: that of Waldspurger is based on a close analysis of distributions on the Lie algebra and their support, whereas the proof (sketched in) [CHLo] is based on motivic integration and model theory.

Comparing the statements of the last four theorems, we conclude that both the Fundamental Lemma for unramified Hecke algebras and the Langlands-Shelstad Transfer Conjecture are valid for all p-adic fields, and for local fields of positive characteristic that is sufficiently large relative to the rank of the group. In the next section we explain how Arthur uses these assertions, and variants thereof, to prove a stable trace formula that can be used to establish functoriality in a variety of situations.

Since parametrization is not available for general p-adic groups, it is not clear what property (II.3.2) means concretely for representations not in the principal series. In practice, Arthur uses the stable trace formula, which incorporates (II.3.1), together with the Fundamental Lemma, which implies (II.3.2) for unramified representations, to *define* local packets for local classical groups in terms of the stable distributions that make (II.3.2) valid in comparison with the stable twisted trace formula for $GL(n)$. Thus local class field theory for classical groups, at present, is derived as a consequence of a global argument based on the stable trace formula.

II.3.13. On Ngô's proof of the Fundamental Lemma.

A full account of Ngô's proof of the Fundamental Lemma is beyond the scope of this article. A good introduction is provided by the article [DT] of Dat and Tuan, and the most difficult technical point is treated in Ngô's short article [N3].[13]

Because the Fundamental Lemma is crucial to the stabilization of the trace formula, it's worth saying a few words about the proof. The local field F of characteristic p is assumed to be the completion at a point v_0 of a curve X over a finite field $\mathbb{F}$. The *Hitchin moduli stack* $\mathcal{M}$ is an algebraic stack over $\mathbb{F}$ classifying G-bundles E over X endowed with a *Higgs field* ϕ, which is roughly sections of the induced $Lie(G)$-bundle over X (via the adjoint representation) with poles along a divisor of sufficiently high degree. The *Hitchin fibration* is the map from $\mathcal{M}$ to an affine space $\mathbb{A}$ that associates to the pair (E, ϕ) a finite collection of invariants of ϕ generalizing the coefficients of the characteristic polynomial of an endomorphism. The points $a \in \mathbb{A}$ define local points a_v for (almost all) points v of X, which can be identified with local stable conjugacy classes in $Lie(g)(k(X)_v)$. The set of points of the Hitchin fiber $\mathcal{M}_a$, modulo an equivalence relation defined geometrically, factors as a product over v of points of a local quotient at v – the quotient of an *affine Springer fiber* by a certain *Picard stack* – and the number of local points can be interpreted as an orbital integral (in the Lie algebra). On the other hand, the number of global points can be counted by using the Grothendieck-Lefschetz trace

[13] See also http://www.time.com/time/specials/packages/article/0,28804,1945379_1944416_1944435,00.html

formula. More subtly, the expressions that occur on the right-hand side of (II.3.8), often called *κ-orbital integrals*, can also be obtained by applying the Grothendieck-Lefschetz trace formula to appropriate perverse sheaves constructed by analyzing the Hitchin fibration. After a series of reductions, the proof of the Fundamental Lemma then amounts to a geometric comparison between the perverse sheaf corresponding to the κ-orbital integral on the right-hand side of (II.3.8), for the Hitchin fibration for G, and the perverse sheaf for the left-hand side of (II.3.8), for the Hitchin fibration for the endoscopic group H. This geometric comparison is quite delicate, not least because the objects in question are stacks rather than schemes, but the rigidity properties of perverse sheaves reduce the comparison in general to the comparison in the most regular situation, where it is more or less elementary.

II.4. Arthur's stabilization of the trace formula, and applications

II.4.1. How to use the stable trace formula.

Endoscopy is an art insofar as it is left to one's discretion to choose the functions f to which one applies the formula (II.2.3), which I reproduce here:

$$T_{disc}(f) = \sum_{(H,s,\xi)\in\mathfrak{E}} i(G,H) ST^H_{disc}(f^H). \tag{II.2.3}$$

(Bear in mind that only the elliptic contribution to the terms on the two sides of this equality have been defined, so the formula as presented here is only a simple approximation to the full stable trace formula proved by Arthur, as we explain below.) However, this art is constrained, primarily by the Fundamental Lemma, and it is this that makes it possible to apply the stable trace formula to Langlands functoriality. Suppose our goal is to prove functorial transfer of L-packets from the endoscopic datum (H,s,ξ) to G, with $\xi : {}^LH \to {}^LG$ the chosen L-homomorphism. One fixes a Langlands parameter ϕ defining an L-packet Π_ϕ on H, and then attempts to find test functions f such that

(II.4.1.1) All terms other than $ST^H_{disc}(f^H)$ vanish, and
(II.4.1.2) If τ is an automorphic representation of H, then $tr\ \tau(f^H) = 0$ unless $\tau \in \Pi_\phi$, in which case $tr\ \tau(f^H) = 1$.

Applying the stable trace formula inductively to the $(H,s,\xi)\in\mathfrak{E}$, one finds that the left-hand side of (II.2.3) $T_{disc}(f) \neq 0$. Writing $T_{disc}(f)$ in its spectral expansion, as in (II.1.5), we find

$$\sum_\pi m_\pi tr\ \pi(f) \neq 0$$

and then we can define the transferred L-packet $\Pi_{\xi_*(\phi)}$ – here we would have $\xi_*(\phi) = \phi\circ\xi$ if Langlands parameters could be defined as genuine homomorphisms – to be the set of π such that $tr\ \pi(f) \neq 0$. The Fundamental Lemma in the form (II.3.6) guarantees that for almost all v such that the elements of Π_ϕ are unramified at v, such a π_v is obtained by functoriality for unramified representations as in (I.2.4). At the remaining places there should be only finitely many possible π_v, which we define to be the local packet $\Pi_{\xi_*(\phi_v)}$, and this definition of $\Pi_{\xi_*(\phi)}$ gives a recipe for local endoscopic transfer.

Although the above sketch is a gross simplification, this strategy works rather well when H is the unique principal endoscopic group, if there is one; it is the basis of Arthur's results on stable transfer from classical groups to $GL(n)$ [A05, A10], as well as the earlier work on base change described in (II.2.4). Of course (II.4.1.2) defines a constraint on test functions on H rather than on G, and it is not always easy to see how to find f satisfying this condition as well as (II.4.1.1). For general H, it is usually impossible to isolate the term ST^H_{disc} by choosing $f \in C_c^\infty(G(\mathbf{A}))$; instead one starts with condition (II.4.1.2) and hopes to be able to compute the contributions $ST^{H'}_{disc}(f^{H'})$. There are interesting cancellations which give rise to the multiplicity formulas conjectured by Arthur; examples will be discussed in §II.5.

II.4.2. Arthur's version of the stable trace formula.

The work of Langlands and Kottwitz in the 1980s had culminated in a full stabilization of the elliptic part of the geometric side of the Arthur-Selberg trace formula, assuming the Transfer Conjecture and Fundamental Lemma. A few years before the proof of the Fundamental Lemma, Arthur had completed stabilization of the trace formula by extending the constructions of Langlands and Kottwitz to the *parabolic* (i.e., non-elliptic) terms. Arthur's strategy involves an elaborate induction that requires simultaneous treatment of the geometric and spectral sides. It also presupposes the Fundamental Lemma as well as a variant adapted to the *weighted orbital integrals*, the distributions Arthur invented to accomodate the parabolic geometric terms. The proof of the "Weighted Fundamental Lemma" for local fields of positive characteristic has recently been completed by Chaudouard and Laumon [CL1, CL2] using the Hitchin fibration as in Ngô's work. The analogue of the theorem on independence of characteristic is proved by Waldspurger in [W09] and by Cluckers, Hales, and Loeser in [CHLo].

Arthur's stable trace formula is an abstract identity of geometric and spectral distributions, where parabolic terms on each side are defined inductively. For this reason it is not practical to state the formula here. We refer the reader for the statements to [A02] and [A05,§§27-29], and instead concentrate on the applications to functoriality.

Arthur proves functoriality for classical groups in [A10] by stabilizing the *twisted trace formula* for $GL(n)$, cf. [A05, (30.4)].[14] As indicated in II.2.4, the stable twisted trace formula is formally analogous to (II.2.3), with $\mathfrak{E}$ the set of twisted endoscopic data. The reader will not be surprised to learn that the proof of the stable twisted trace formula requires yet another round of Fundamental Lemmas. Nor will the reader be surprised to see the names of Waldspurger and Ngô in connection with the solution of the attendant problems. The sequence of steps leading to the solution begins with Waldspurger's reduction in [W08] of the Twisted Fundamental Lemma (relating a twisted orbital integral on the Lie algebra of G to a stable orbital integral on the Lie algebra of the twisted endoscopic group H) to an identity of stable orbital integrals on two auxiliary Lie algebras, with no twisting in sight. This identity, which Waldspurger called the *Non Standard Fundamental Lemma* was proved by Ngô, in the case of positive characteristic, together with the proof of the Fundamental Lemma II.3.8 [N2].[15]

[14]In his 2008 lectures at Banff, Arthur indicated that, although the necessary Fundamental Lemmas are available, this stabilization had not yet been carried out in detail along the lines of [A02]. Many of the preliminary steps are now available, but specialists expect that full stabilization will require sustained efforts by a number of people.

[15]The attentive reader may have noticed that I have not stated the analogues of Theorems II.3.10 and II.3.12 for non-standard endoscopy. In the interests of full disclosure, it should be

The statements of Arthur's main applications of the stable twisted trace formula have been available for some time [A05, §30]; the proofs will appear in [A10]. For the results I am about to discuss, I think it is safe to assume that the strategy is roughly the one sketched in (II.4.1), applied to the twisted group $\widetilde{GL}(n) := GL(n) \rtimes \Theta$ of §II.2.4. When $n = 2k+1$ is odd, H is the unique principal endoscopic group $H_k = Sp(2k)$, but taken infinitely often depending on the choice of L-embedding ξ_η; when $n = 2k$ is even, H is either the principal endoscopic group $SO(2k+1)$ or a principal endoscopic group of the form $SO(2k, \eta)$ (see [A05, §30] for this notation). Distinguishing among the representations arising from the various members of this infinite collection of principal endoscopic groups is a supplementary problem that we will not attempt to address here.

II.4.3. Applications to the local Langlands correspondence.

In the proofs of functoriality based on a comparison of trace formulas, the local and global questions are not neatly separated. In order to state Arthur's results, however, it is more convenient to follow Arthur's exposition in [A05, A08] and present the local results first.

The comparison of trace formulas is not only used to establish transfer from the endoscopic groups to $GL(n)$ but also to characterize the image in the discrete (and not necessarily discrete) spectrum of $GL(n)$. This is used in the reverse direction to parametrize local and global representations of the classical groups in terms of representations of $GL(n)$. In the local situation the irreducible admissible representations of $GL(n, F)$ are classified by the local Langlands correspondence, and in this way local parameters (I.2.2) classify packets on the classical group G.

In this section F is a non-archimedean local field of characteristic zero. We follow Arthur and let N be a positive integer,

$$\phi : \mathcal{L}_F = W_F \times SU(2) \to GL(N, \mathbb{C}) \times W_F$$

be a local Langlands parameter as in (I.2.2), and assume ϕ is *self-dual*. Then $Im(\phi)$ stabilizes a decomposition

$$\mathbb{C}^N = V_s(\phi) \oplus V_o(\phi) \oplus V_{split}(\phi) = V_s \oplus V_o \oplus V_{split}$$

where V_s (resp. V_o) is endowed with a non-degenerate symplectic (resp. symmetric) form, fixed by $Im(\phi)$, and $V_{split} = W \oplus Hom(W, \mathbb{C})$ with its natural duality. We assume $V_{split} = 0$ (this is the elliptic condition).

Definition II.4.3.1 [A05,§30]. *Let ϕ be a parameter as above, and write*

$$V_s(\phi) = \oplus_{i \in I_s} V(\phi_{i,s})^{\ell_i}; V_o(\phi) = \oplus_{i' \in I_o} V(\phi_{i',o})^{\ell_{i'}}; V_{split} = \oplus_{j \in J}[W(\phi_j) \oplus W(\phi_j)^\vee]^{\ell_j}$$

where each summand is an irreducible representation of $\mathcal{L}_F$, the ℓ_* are multiplicities, and $W(\phi_j)^\vee$ denotes the dual of $W(\phi_j)$.

(i) *We say ϕ is discrete if $V_{split} = 0$ and all $\ell_i = \ell_{i'} = 1$.*

mentioned here that the proofs of these analogues have not been written down, but Waldspurger assures me that they are identical to the proofs for standard endoscopy. Nor has anyone written down the analogue of Hales deduction of Theorem II.3.6 from Theorem II.3.8 for twisted endoscopy. The reader may prefer to consider the applications described below conditional on the complete verification of these stray endoscopic identities.

(ii) *We say ϕ is primitive if $\phi(\mathcal{L}_F)$ acts irreducibly on V. In particular, either $V = V_s$ or $V = V_o$.*
(iii) *For any twisted endoscopic group G of $GL(N,F)$, let $\tilde{\Phi}(G)$ denote the set of (equivalence classes of)*[16] *self-dual ϕ that factor through the image of ${}^LG$ in ${}^LGL(N)$.*

A primitive ϕ has image in $Sp(N,\mathbb{C}) \times W_F$ (resp. $O(N,\mathbb{C}) \rtimes W_F$ where there may be a twisted action if N is even) if $V = V_s$ (resp. $V = V_o$). In general, ϕ defines a self-dual representation π_ϕ of $GL(N,F)$ by the local Langlands correspondence; π_ϕ is (essentially) square-integrable if and only if ϕ is primitive. (Note that primitive ϕ correspond to *indecomposable* Weil-Deligne parameters.)

In general, let $\phi \in \tilde{\Phi}(G)$, and define $S_\phi(G)$ to be the centralizer in ${}^LG$ of $Im(\phi)$, $\mathcal{S}_\phi(G) = S_\phi(G)/(S^0_\phi(G) \cdot Z(\hat{G}^{\Gamma_F}))$. If ϕ is primitive then $\mathcal{S}_\phi(G)$ is trivial; in general it is a group of the form $(\mathbb{Z}/2\mathbb{Z})^r$ where $r = |I_s|+|I_o|$ or $r = |I_s|+|I_o|-1$. The order of the group $\mathcal{S}_\phi(G)$ is meant to determine the number of representations of G in the packet attached to ϕ. In particular, Arthur proves that a primitive π_ϕ descends to an irreducible representation of a unique principal endoscopic group for $GL(N,F)$, which is – almost – determined as follows. We let $\eta_\phi = \det \circ \phi : W_F \to \mathbb{C}^\times$; by duality η_ϕ is a character of order 1 or 2. If η_ϕ is non-trivial, it determines a unique principal endoscopic pair (H_k, ξ_{η_ϕ}) (if $N = 2k+1$ is odd) or $SO(2k, \eta_\phi)$ (if $N = 2k$ is even), and this is the group to which π_ϕ descends. This remains true even if $\eta_\phi = 1$ for N odd. However, if $N = 2k$ is even and $\eta_\phi = 1$, then the two possibilities H_k and $SO(2k,1)$ remain, and have to be distinguished by an L-function criterion.

Now assume $N = 2n$ is even and $\phi \in \tilde{\Phi}(H_n)$; thus ϕ takes image in $Sp(N,\mathbb{C}) \times W_F$ and $\eta_\phi = 1$. On the other hand, π_ϕ extends canonically[17] to the twisted group $GL(N,F) \rtimes \Theta$ and thus determines a canonical twisted character $f \mapsto tr\ \tilde{\pi}_\phi(f)$ for f a test function on $GL(N,F) \rtimes \Theta$.

For any p-adic reductive group G, let $\Pi_{fin,unit}(G)$ denote the set of formal finite linear combinations $\sum a(\pi)\pi$ of irreducible admissible unitary representations of G, with $a(\pi) \geq 0$ in $\mathbb{Z}$. Recall that a unitary representation of G is called *tempered* if it belongs to the support of Plancherel measure, or equivalently if its matrix coefficients are in $L^{2+\varepsilon}(G)$ for all $\varepsilon > 0$. Tempered representations are the building blocks of the Langlands classification (which, for p-adic groups, is significantly coarser than the local Langlands correspondence.) The following theorem includes the local Langlands correspondence for tempered representations of G.

Theorem II.4.3.2 (Arthur). *Let G be a principal twisted endoscopic group for $GL(N)$ and let $\phi \in \tilde{\Phi}(G)$.*

(a) Under the above hypotheses, there is a stable linear form on test functions on G:

$$h \mapsto D_{\phi,G}(h)$$

that is the transfer of the twisted character $tr\ \tilde{\pi}_\phi$ in the sense that

$$D_{\phi,G}(f^G) = tr\ \tilde{\pi}_\phi(f)$$

[16] Arthur uses a notion of equivalence slightly coarser than $\hat{G}$-conjugacy; roughly, two parameters for $SO(2k)$ are equivalent if they are conjugate under $O(2k)$. See the discussion in [A05] following (30.10).

[17] ... provided it is generic; otherwise an induction step is needed at this point.

for any test function f on $GL(N,F) \rtimes \Theta$. Here f^G is the (twisted) transfer of f to a test function on G.

(b) There is a finite subset Π_ϕ of $\Pi_{fin,unit}(G)$ and a bijection

$$\Pi_\phi \mapsto \hat{\mathcal{S}}_\phi(G) = Hom(\mathcal{S}_\phi(G), U(1));\ \pi \mapsto <\bullet, \pi>$$

that satisfies the following condition. For any semisimple $s' \in S_\phi(G)$, with image $x \in \mathcal{S}_\phi(G)$, and with (H', ϕ') the preimage of (ϕ, s') in the sense described in (II.4.3.3) below,

$$D_{\phi',H'}(f^{H'}) = \sum_{\pi \in \Pi_\phi} <x, \pi> tr\ \pi(f)$$

for any test function f on G.

(c) All elements of Π_ϕ are tempered.

(d) Finally, if $\phi \neq \phi'$ are two inequivalent elements of $\tilde{\Phi}_{prim}(G)$ then $\Pi_\phi \cap \Pi_{\phi'} = \emptyset$.

II.4.3.3. In (b) the parameter ϕ is assumed to factor through a collection of subgroups of the L-group of G, parametrized by the elements $s' \in S_\phi(G)$. Each such s determines an endoscopic group H' for G and the factorization implies that ϕ comes from a parameter ϕ' with values in the L-group of H'. This pair (H', ϕ') is the preimage to which (b) refers. The function $f^{H'}$ is then the endoscopic transfer of f to H', and $\bullet \mapsto D_{\phi',H'}(\bullet)$ is a stable linear form on test functions on H' analogous to the one whose existence is asserted in (a). Indeed, the statement of (a more general version of) this theorem in [A05] is preceded by an induction hypothesis that presupposes that the stable distributions of (a) have already been constructed for all possible endoscopic groups for G. When $G = H_n$, we have seen in §II.2 that the elliptic endoscopic groups for G are all odd orthogonal groups, so this is reasonable, but in general one needs to treat all classical groups simultaneously in order to make sense of this induction.

The coefficients $<\kappa, \pi>$ in (b) are examples of the "spectral transfer factors" mentioned in the discussion preceding (II.3.6). The formula in (b) is a version of the expected endoscopic character identity. The group $\mathcal{S}_\phi(G)$ is always a finite 2-group, so the coefficients $<\kappa, \pi>$ are always signs. This is also true for the unitary groups discussed below.

A Langlands parameter ϕ for G is an *Arthur parameter*

(II.4.3.4) $$\psi : \mathcal{L}_F \times SU(2) = W_F \times SU(2) \times SU(2) \to {}^LG$$

that is trivial on the second factor $SU(2)$. We define $\tilde{\Psi}_{prim}(G)$ to be the set of equivalence classes (see footnote 16) of primitive Arthur parameters. The analogue of Theorem II.4.3.2 for $\tilde{\Psi}_{prim}(G)$ is stated as Theorem 30.1 of [A05]. The character formula (b) involves an additional sign arising from the action of the second $SU(2)$. Conditions (c) and (d) both fail for general Arthur parameters; indeed, the non-triviality of the map on the second $SU(2)$ guarantees that the *Arthur packet* Π_ψ includes non-tempered representations. We refer the reader to [A05] and [A08] (especially the lecture of August 19, 2008) for more details.

II.4.3.5. Remark on characterization of the local correspondence. The local correspondence Arthur constructs for classical groups is characterized by the character relation (Theorem II.4.3.2 (b) and its analogues for other groups) which is thus linked to the local Langlands correspondence [HT,He] for $GL(N)$. The latter correspondence is in turn characterized by a short list of axioms, the most important of which are compatibility (via abelian local class field theory) with twists by abelian characters, and preservation of local L and ε- factors. Local L- and ε- factors more general than those that arise from transfer to $GL(N)$ can be constructed directly in some cases for classical groups. As far as I know, trace formula techniques cannot be used to link the local correspondence for classical groups to these constructions without passing through $GL(N)$.

Since the local packets for classical groups are defined formally in terms of irreducible representations of $GL(N)$, local transfer from classical groups to $GL(N)$ can only be understood tautologically. This step is necessary, however, in order to define global functorial transfer for classical groups.

II.4.3.6. Structure of local Arthur packets, after Moeglin. Alone and in collaboration with M. Tadić and J.-L. Waldspurger, Colette Moeglin has established an extensive theory of the structure of representations of p-adic classical groups. Most relevant to the topic of the present paper are her recent articles [M09,M09a] on the structure of local Arthur packets. Let G be a classical group, ψ an Arthur parameter for G, and $D_{\psi,G}$ the stable distribution constructed by Arthur (of which D_{ϕ,H_n} of (II.4.3.2) is a special case). Now $D_{\psi,G}$ can be written as a finite linear combination of characters of irreducible admissible representations with integer coefficients. Assuming the coefficients are all equal to 1 when ψ is a *discrete series* packet (in particular, $\psi = \phi$ is a Langlands parameter), Moeglin shows that the same is true for any tempered parameter, and moreover the coefficients are always ± 1 for any Arthur packet, with the sign determined explicitly. In other words, if the discrete series for G have multiplicity 1 in their local packets, then any unitary representation occurs with coefficient ± 1 in some stable integral linear combination of characters. In private communication, Moeglin has sketched a proof she attributes to Arthur of the multiplicity 1 assertion for discrete series.

Using Labesse's results on the simple trace formula and base change, described in §II.5 below, Moeglin has also constructed stable packets for p-adic unitary groups [M07].

II.4.4. Global transfer for classical groups.

Global transfer is formally analogous to local transfer, and our treatment here will be brief. We have no global analogue of the local Langlands groups, so, as explained in §I.3, we construct formal Arthur parameters (I.3.4) using cuspidal automorphic representations of $GL(a_i)$. In this section we first consider descent to a classical group G of cuspidal automorphic representations Π of $GL(N)$; we then explain Arthur's multiplicity formula for the discrete spectrum of G.

Let Π be a self-dual cuspidal automorphic representation of $GL(N)$. The central character $\eta = \eta_\Pi$ is a Hecke character of order 1 or 2. As in the discussion above, a non-trivial η determines a unique principal twisted endoscopic datum (G_η, s, ξ_η). If N is even and η is trivial, then there are two possibilities that are distinguished as follows. It is known that the Rankin-Selberg L-function $L(s, \Pi \otimes \Pi^\vee$ has a simple pole at $s = 1$ [Sh,JPSS]. But $\Pi^\vee \simeq \Pi$, and

$$L(s, \Pi \otimes \Pi^\vee) = L(s, \Pi, Sym^2) \cdot L(s, \Pi, \wedge^2)$$

Exactly one of the two factors on the right has a pole. If it is the first factor then the relevant twisted endoscopic group G is the split $SO(N)$, whereas if it is the second factor then $G = SO(N+1)$ [A05, Theorem 30.3].

In this way Π determines a unique twisted endoscopic datum (G, s, ξ). We let ϕ_G be a symbol denoting the formal notation for the Langlands parameter of the packet on G that transfers to Π on G. It follows from Arthur's constructions that, for every place v of F, the local Langlands parameter $\phi_v = \phi_{\Pi_v}$ of the component Π_v of Π belongs to $\tilde{\Phi}(G_v)$, where $G_v = G(F_v)$. Theorem II.4.3.2 thus determines a local packet Π_{ϕ_v} for each v. Now define

$$\Pi_{\phi_G} = \{\bigotimes_v \pi_v \ : \ \pi_v \in \Pi_{\phi_v}, < \bullet, \pi_v > \equiv 1 \text{ for almost all } v\}. \tag{II.4.4.1}$$

In this case Arthur proves that every element of Π_{ϕ_G} occurs with multiplicity 1 or 2 (the latter only when $G = SO(2k)$) in the discrete automorphic spectrum of G. This is the *stable* part of the discrete automorphic spectrum of G.

More generally, let $N = \sum_{i=1}^{r} N_i$ and let Π_i be a self-dual cuspidal automorphic representation of $GL(N_i)$. In [A05, pp. 242-243], Arthur gives a recipe for deciding when the formal parameter (Π_i) for $\prod_i GL(N_i)$ defines a parameter ϕ_G for the chosen endoscopic datum (G, s, ξ). We again obtain local packets Π_{ϕ_v} and define Π_{ϕ_G} by the recipe (II.4.4.1). Moreover, Arthur defines groups S_{ϕ_G} and $\mathcal{S}_{\phi_G}$ analogous to the local versions introduced above [A05, (30.11), (30.12)], together with maps

$$\mathcal{S}_{\phi_G} \to \mathcal{S}_{\phi_v}$$

for all places v of F. Thus every character $\kappa \in \hat{\mathcal{S}}_{\phi_v}$ restricts to a character of $\mathcal{S}_{\phi_G}$.

Here is an abbreviated version of Theorem 30.2 of [A05].

Theorem II.4.4.2 (Arthur). *For every $\pi \in \Pi_{\phi_G}$, define a character*

$$< \bullet, \pi > = \prod_v < \bullet, \pi_v > : \mathcal{S}_{\phi_G} \to U(1).$$

Then π occurs as a direct summand of $\mathcal{A}^2(G)$ with non-zero multiplicity if and only if $< \bullet, \pi > = 1$. In that case, the multiplicity is either 1 or 2, the latter only occurring if $\hat{G} = SO(N, \mathbb{C})$ with N even.

Remark. The multiplicity result depends in an essential way on the Jacquet-Shalika classification theorem [JS], a generalization of the strong multiplicity one theorem for cuspidal automorphic representations of $GL(N)$ and the precise automorphic analogue of the Chebotarev density theorem for Galois representations. In particular, an automorphic representation Π of $GL(N)$ is isomorphic to its contragredient as abstract representation if and only if it is **equal** to its dual as a space of automorphic forms on the adèle group of $GL(N)$.

In the stable case, when $r = 1$, $\mathcal{S}_{\phi_G} = \{1\}$. When $\mathcal{S}_{\phi_G}$ is non-trivial, the condition that $< \bullet, \pi > = 1$ is not vacuous, and certain elements of Π_{ϕ_G} do not occur as automorphic representations. Those that do occur are often called *endoscopic.* Thus to a first approximation, endoscopic automorphic representations correspond to reducible Langlands parameters.

The parameters ϕ_G correspond to expressions (I.3.4) with all $b_i = 1$. More generally, Arthur defines parameters ψ_G for every expression (I.3.4) attached to

G; the b_i corresponds to the dimension of a representation of the group $SU(2)$ we have already seen as the *second* $SU(2)$ in our discussion of local Arthur parameters in (II.4.3.4). In particular, ψ_G determines local Arthur parameters ψ_v. The generalization to Arthur parameters of Theorem II.4.3.2 defines local packets Π_{ψ_v} and a global packet Π_{ψ_G} with the formula (II.4.4.1). Theorem 30.2 of [A05] includes a multiplicity formula for $\pi \in \Pi_{\psi_G}$. The multiplicity of π is positive provided $< \bullet, \pi >= \kappa_\psi$ for an explicit character κ_ψ constructed out of symplectic root numbers.

Finally, Arthur proves

Theorem II.4.4.3. *Every discrete automorphic representation of G arises from a formal parameter (I.3.4) in this way.*

Theorems II.4.4.2 (including the version for Arthur parameters) and II.4.4.3 imply formally the existence of functorial transfer from discrete automorphic representations of G to automorphic representations of $GL(N)$, and determine the image of this transfer. In other words, these theorems completely solve the problem of functoriality in these cases.

II.5. Base change in a simple situation

The results of the previous section make full use of the stabilized trace formula of [A02,A05], and there is no way to prove functoriality in general without a complete understanding of the parabolic terms. The lesson of the Paris book project on the trace formula and Galois representations [Book1,Book2] is that, for many, if not most, applications to arithmetic, it suffices to use the simple version of Arthur's stable trace formula, proved in [A88], in which the terms on left-hand (resp. right-hand) side of (II.2.3) are genuine sums of orbital integrals (resp. stable orbital integrals) of elliptic elements. The results described in the present section are all applications of the simple trace formula. Of course one must bear in mind that the simple trace formula is only obtained as a consequence of the full trace formula; nevertheless, once it is available, it is indeed much simpler to use.

Let $U = U_n$ relative to a local or global quadratic extension F'/F, as in §I.1, $G = R_{F'/F}U_{F'} \xrightarrow{\sim} {}^LGL(n)_{F'}$, and consider the L-homomorphism $BC : {}^LU \to {}^LG$ defined in [Mi, 4.1.2] (standard base change). This is a case of base change, for which the relevant Fundamental Lemmas were proved long ago in [K86a] (for the analogue of II.3.8) and [C90,L90] (for the analogue of II.3.6; see also [CL]). The Transfer Conjecture was proved by Waldspurger. For global F there is thus the basis for the comparison between the twisted trace on $G \rtimes \Theta$, where $\Theta = Gal(F'/F)$ as in II.2.4, and the stable traces $ST_{disc}^{U^*}$ on the quasi-split inner form U^* of U and the remaining twisted endoscopic groups of G. This alone does not suffice to relate automorphic representations of U to Θ-stable automorphic representations of G. The former are measured by the trace T_{disc} on U, which in the best of circumstances satisfies (II.2.3), a sum in which $ST_{disc}^{U^*}$ appears along with the endoscopic contributions $ST_{disc}^{U_{k,j}^*}$ with $U_{k,j}^* = U_k^* \times U_j^*$, $k + j = n$, $k \leq j$ (cf. (II.2.2.3)).

We assume F to be a totally real number field, F' a totally imaginary quadratic extension. Simplification of the comparison begins with the following observation due to Labesse. Let $f = f_\infty \otimes f_f$ be a test function on $G(\mathbf{A}) \rtimes \Theta$, and assume $f_\infty = Lef_W$ is a *Lefschetz function* for base change, attached to a finite-dimensional representation W of $R_{F/\mathbb{Q}}U$.[18] This is the analogue for the twisted trace formula of the Euler-Poincaré functions introduced in II.1.7, and has the property that it only has non-zero trace on representations that are cohomological for the base change $W \otimes W^\theta$ of W to $G(F' \otimes_{\mathbb{Q}} \mathbb{R})$.[19]

Theorem II.5.1 (Labesse [Lab09, Theorem 4.12]). *Let $f = Lef_W \otimes f_f$ be a test function on $G(\mathbf{A}) \rtimes \Theta$, and let f^{U^*} be an associated function on $U^*(\mathbf{A})$. Then*

$$T_e^{G \rtimes \Theta}(f) = ST_e(f^{U^*})$$

where the subscript $_e$ denotes the sum over elliptic conjugacy classes.

In other words, U^* is the only twisted endoscopic group for $G \rtimes \Theta$ to contribute. Up to a constant [Lab09, Lemme 4.4] we have the relation

$$Lef_W^{U^*} = EP_W$$

for the stable transfer of $f_{\infty,W}$. In particular,

$$T_{disc}^{G \rtimes \Theta}(f) = \sum_{\Pi \xrightarrow{\sim} \Pi^\theta} tr\ \Pi(f)$$

where Π runs over θ-invariant representations that are *cohomological* for $W \otimes W^\theta$, while

$$T_{disc}^{U^*}(f^{U^*}) = \sum_{\pi} m(\pi) tr\ \pi(f^{U^*})$$

where π runs over representations that are cohomological for W.

Now suppose $[F : \mathbb{Q}] \geq 2$. Then the choice of test functions at infinity allows Labesse to apply Arthur's simple trace formula and (under additional regularity hypotheses) to replace the left-hand side of (II.5.1) by $T_{disc}^{G \rtimes \Theta}(f)$ [Lab09, Props. 3.9, 3.10]. Applying II.5.1 to each of the terms $ST_{disc}^{U_{k,j}^*}$ of the stabilization (II.2.3) of the *elliptic part* of the trace formula for our original unitary group U, we then find

Theorem II.5.2 [Lab09, Theorem 5.1]). *Assume $[F : \mathbb{Q}] \geq 2$ and $\varphi = \varphi_\infty \otimes \varphi_f$ is a test function on $U(\mathbf{A})$ such that $\varphi_\infty = EP_W$, or more generally is a pseudo-coefficient of a discrete series representation of $U(F \otimes \mathbb{R})$ that is cohomological for W. Then*

$$T_{disc}^U(\varphi) = \sum_{k+j=n, k \leq j} i(U, U_k^* \times U_j^*) T_{disc}^{G_{k,j} \rtimes \Theta}(\varphi^{G_{k,j}}).$$

Here $G_{k,j} = R_{F'/F} U_{k,j}^*$, the action of Θ is as in (I.1.2), and $\varphi^{G_{k,j}}$ is a test function on $G_{k,j} \rtimes \Theta$ [Lab09, Prop. 4.12] whose transfer to the twisted endoscopic group $U_{k,j}^*$ equals $\varphi^{U_{k,j}^*}$.

[18]For clarity, I point out that W can be written $\otimes_{v \mid \infty} W_v$ where each W_v is an irreducible representation of $U(F_v)$. Thus Lef_W also factors as a product of test functions over real places of F.

[19]References to [Lab09] are based on the version available online in 2009. The numbering of statements in the published version may be slightly different.

In this way the cohomological part of the trace T^U_{disc} is expressed entirely in terms of the trace formulas for products of general linear groups. The subscript $_{disc}$ on the right-hand side is deceptive however, because a θ-invariant representation of $G_{k,j}$ can contribute discretely to the twisted trace formula even if it belongs to the continuous automorphic spectrum of $G_{k,j}$. These contributions can be determined, however, and the following results are proved in [Book1] on the basis of the above theorems, making use of the stabilization [LN,W06] of the elliptic part of the trace formula, but without referring to Arthur's complete stabilization [A02].

Remark. As the referee pointed out, Labesse's proof of Theorem II.5.2 makes crucial use of the fact [Lab09, Lemme 4.1] that the local transfer map $f \circ\!\!\to f^{U^*}$ is *surjective* at non-archimedean places. This is not true in general for twisted endoscopy.

Let Π be a cuspidal automorphic representation of G such that $\Pi \xrightarrow{\sim} \Pi^\theta$ and Π_∞ is cohomological for some $W \otimes W^\theta$. We begin by making a series of simplifying hypotheses.

Simplifying Hypotheses II.5.3.

(1) *F'/F is unramified at all finite places (in particular $[F:\mathbb{Q}] \geq 2$).*
(2) *Π_v is spherical (unramified) at all non-split non-archimedean places v of F'.*
(3) *The degree $d = [F:\mathbb{Q}]$ is even.*
(4) *At all non-archimedean places the group $U(F_v)$ is quasi-split.*

In the arithmetic applications discussed below one easily reduces to the situation of II.5.3. By (1) all places are either split finite, inert finite, or real. We apply the formula in Theorem II.5.2. Write

$$\varphi = \otimes' \varphi_v,$$

where $\otimes'$ means that at almost all places φ_v is the identity element in the spherical Hecke algebra. We have already assumed $\otimes_{v \mid \infty} \varphi_v = Lef_W$; as above (but with the matching going in the opposite direction) this matches $\varphi_\infty = EP_W$. At inert v Π_v is unramified, so we can take φ_v in the spherical Hecke algebra at v for the hyperspecial maximal compact subgroup, and each $\varphi_v^{G_{k,j}}$ is determined by (two applications of) the Fundamental Lemma II.3.6. Finally, at split v it is easy to choose explicit matching functions φ_v and $\varphi_v^{G_{k,j}}$ (see [Lab99,§3.4]). Since Π is assumed cuspidal, the Jacquet-Shalika classification theorem for representations of $GL(n)$ [JS] implies that we can choose a finite linear combination $f = \sum_j \varphi_j$, with φ_j as above, with property (II.4.1.1), with $H = GL(n) = G_{0,n}$; in other words only the term corresponding to $GL(n)$ survives on the right-hand side. Moreover, II.5.3 implies f can also be chosen to isolate the chosen representation Π (cf. II.4.1.2). On the other hand, it follows from the explicit nature of these functions and from the simplifying hypotheses that the right hand side of the equation in II.5.2 is of the form

$$T^U_{disc}(f) = \sum_{\pi = \pi_\infty \otimes \pi_f} m(\pi)\ tr\ \pi(f)$$

where π_∞ is cohomological for W and $\pi_f = \otimes' \pi_v$ is the *unique* representation such that Π_v is the functorial transfer of π_v locally everywhere for the L-homomorphism denoted BC above. (This uniqueness makes strong use of II.5.3 (2) and (4), in particular; it implies that the local L-packet on $U(F_v)$ that transfers to the Langlands parameter associated to Π_v contains a single representation unramified with respect to the hyperspecial maximal compact.) The final result is

Theorem II.5.4 [Lab09, Théorème 5.4, Corollaire 5.3]. *Assume Hypotheses II.5.3 and let* π_f *be the unique representation of* $U(\mathbf{A}^f)$ *that transfers to* Π_f*. Let* π_∞ *be any discrete series representation of* $U_\infty := U(F \otimes \mathbb{R})$ *that is cohomological for* W*. Then* $\pi_\infty \otimes \pi_f$ *occurs with positive multiplicity in the discrete automorphic spectrum of* U*. The number of such* π_∞ *is* $\prod_{v \mid \infty} \binom{n}{r_v}$*, where* $U(F_v)$ *has signature* $(r_v, n - r_v)$*. Moreover, if the highest weight of* W *is regular, or if* U_∞ *is a product of compact unitary groups, then each such* $\pi_\infty \otimes \pi_f$ *occurs with multiplicity* 1.

Conversely, if the highest weight of W *is regular, or if* U_∞ *is a product of compact unitary groups, and if* π *is an automorphic representation of* U *that is cohomological for* W *and is spherical at all non-split non-archimedean places* v *of* F*, then* π *admits a base change* Π *to* $GL(n)$ *that is cohomological, stable under* θ*, and satisfies II.5.3 (2).*

It goes without saying that this result is compatible with what is predicted by Arthur's multiplicity conjecture for this class of representations. The base change in the second part of the theorem is *not necessarily cuspidal*; the possibilities are indicated in [Lab03, Cor. 5.3].

In order to construct Galois representations, we also need to consider the contribution of the term $T_{disc}^{G_{1,n-1} \rtimes \Theta}$ to T_{disc}^U. We need two specific classes of U.

Lemma II.5.5. *Hypotheses II.5.3 remain in force. (a) There exists an* n*-dimensional hermitian space* V_0/F' *relative to the extension* F'/F *such that the unitary group* $U_0 = U(V_0)$ *satisfies*

- *For all finite places* v*,* $U_0(F_v)$ *is quasi-split and splits over an unramified extension of* F_v*; in particular,* $U_0(F_v)$ *contains a hyperspecial maximal compact subgroup.*
- *For all real places* v*,* $U_0(F_v)$ *is compact.*

Moreover, U_0 *is unique up to isomorphism.*

(b) Suppose n *is even. There exists a hermitian space* V_1/F' *of dimension* $n+1$ *relative to the extension* F'/F *such that, letting* $U_1 = U(V_1)$*,*

- *For all finite places* v *of* F*,* $U_1(F_v)$ *is quasi-split and splits over an unramified extension of* F_v*; in particular,* $U_1(F_v)$ *contains a hyperspecial maximal compact subgroup.*
- *For all real places* v *of* F *, with the exception of one place* v_0*,* $U_1(F_v)$ *is compact;* $U_1(F_{v_0}) \xrightarrow{\sim} U_1(1, n)$*.*

The proof is a standard application of Galois cohomology, cf. [C91]. Everything we will need about automorphic forms on U_0 is contained in Theorem II.5.4. As for U_1, it is anisotropic and we have seen in II.5.4 that there are $n+1$ discrete series representations of $U_{1,\infty}$ cohomological for a fixed irreducible finite-dimensional representation W_1.

Theorem II.5.6 [CHL1, Theorem 4.6]. *Let* Π *be a cuspidal automorphic representation of* $GL(n)$ *satisfying* $\Pi \xrightarrow{\sim} \Pi^\theta$*, cohomological for* W*. Assume the highest* *weight of W is sufficiently regular (regularity is more than enough). Then there is a cohomological character χ of $GL(1)_{F'}$ that is a stable base change from $U(1)$ (it is trivial on the idèles of F), a finite-dimensional irreducible representation W_1 of $GL(n+1)$ (depending on W and the infinity type of χ), and an L-packet $\pi(\Pi,\chi)$ of automorphic representations $\pi = \pi_\infty \otimes \pi_f$ of U_1 such that*

1. $BC(\pi) = \Pi \boxplus \chi$ *(the Langlands sum, whose standard L-function satisfies $L(s, \Pi \boxplus \chi) = L(s,\Pi)L(s,\chi)$)*
2. π_∞ *is discrete series and cohomological for W_1*
3. *The factor π_f is uniquely determined and denoted $\pi_f(\Pi_f, \chi_f)$.*
4. *Of the possible $n+1$ $\pi_\infty \otimes \pi_f(\Pi_f,\chi_f)$ with base change $\Pi \boxplus \chi$, exactly n occur, each with multiplicity 1.*

Each discrete series that occurs contributes a one-dimensional component to the cohomology. The next section explains the relation of (II.5.4) and (II.5.6) with the construction of a compatible system of ℓ-adic Galois representations attached to Π.

On the regularity hypothesis. The regularity hypothesis of (II.5.4) and (II.5.6) is called *strong regularity* in [CHL2, 1.2.3]. The result holds under a weaker regularity hypothesis but the proof uses an argument of Kottwitz [K92a] based on the purity of Frobenius eigenvalues that requires the results of [CHL2] as well as [CHL1]. The article [Shin] includes a version of this argument. Regularity is in fact irrelevant to (II.5.4) but some hypothesis is needed in II.5.6 even when $n = 2$ [BR].

On the simplifying hypotheses. In the final section of [Lab09], Labesse drops the simplifying hypotheses but admits the full strength of Arthur's stabilization of the trace formula [A02] as well as Moeglin's work on local L-packets for unitary groups [M07]. With the help of these results (whose proofs are well beyond the scope of [Book1]), Labesse is able to prove base change and descent results for general automorphic representations of U of discrete series type at real places, and to show that the local base change at ramified places is compatible with Moeglin's definition of local L-packets.

Part III. Arithmetic applications of the stable trace formula

III.1. Shimura varieties and Galois representations

III.1.1. General considerations.

We have already encountered locally symmetric varieties in §II.1.7. Shimura varieties tend not to be attached to anisotropic groups, so we modify the notation slightly. We assume G to be a reductive group over $\mathbb{Q}$ whose center Z has the property that

$$\mathbb{Q}-\text{rank } Z \;=\; \mathbb{R}-\text{rank } Z \tag{III.1.1}$$

Let $K'_\infty \subset G(\mathbb{R})$ be a maximal compact subgroup, $K_\infty = K'_\infty Z(\mathbb{R})$. With this new definition of K_∞, define $S(G)$ by (II.1.7.1); then (III.1.1) implies that

$$S(G) = \varprojlim_{K_f} {}_{K_f}S(G)$$

with ${}_{K_f}S(G)$ as in II.1.7 and K_f runs over open compact subgroups of $G(\mathbf{A}^f)$. Each ${}_{K_f}S(G)$ is isomorphic to a finite union of quotients of $X_G := G(\mathbb{R})/K_\infty$ by discrete arithmetic subgroups of $G(\mathbb{Q})$. When the pair (G, X_G) is of hermitian type, $S(G)$ is a *Shimura variety*, and each ${}_{K_f}S(G)$ is a quasiprojective variety over $\mathbb{C}$. In Deligne's formalism [D71], the datum X_G can be identified, generally in more than one way, with a $G(\mathbb{R})$-conjugacy class of homomorphisms $R_{\mathbb{C}/\mathbb{R}}(\mathbb{G}_m)_{\mathbb{C}} \to G_{\mathbb{R}}$. Fixing one such identification determines a *reflex field* $E(G, X_G)$, finite over $\mathbb{Q}$, and a *canonical model* of $S(G)$ as a proalgebraic variety with $G(\mathbf{A}^f)$-action over $E(G, X_G)$. We will write $E(G)$ for simplicity.

The arithmetic theory of Shimura varieties is of course due to Shimura, who introduced the *PEL types* of particular importance in applications ([Sh66] and the many articles mentioned there) as well as proving the existence of canonical models for a larger class of varieties and formulating the general conjecture (summarized in [Sh70]). The variety $S(G)$ is of PEL type if it can be identified with the moduli space of abelian varieties with specified polarization, endomorphisms, and level structures, with all data satisfying natural compatibilities. The reformulation of these conditions in the language of representable functors, adapted to moduli spaces in mixed characteristic, is due to Kottwitz [K92b]. For a recent comprehensive introduction to the theory of Shimura varieties, see [Ml05].

Let (ρ, W) be a finite-dimensional algebraic representation of G, and define $\tilde{W}_{K_f}$ as in II.1.7, $\tilde{W}$ the projective limit over $S(G)$. This is a local system with coefficients in a finite extension $E(W)$ of $E(G)$ in general, and so for each i the (topological) cohomology $H^i(S(G), \tilde{W})$ is naturally a vector space over $E(W)$ with an algebraic action of $G(\mathbf{A}^f)$. For each rational prime ℓ and each prime divisor $\lambda \mid \ell$ of $E(W)$ the tensor product $\tilde{W}_\lambda := \tilde{W} \otimes_{E(W)} E(W)_\lambda$ defines an étale local system over $S(G)$ in $E(W)_\lambda$-vector spaces, and thus

$$H^i_\lambda(S(G), \tilde{W}) := H^i(S(G), \tilde{W}_\lambda)$$

is an $E(W)_\lambda$-vector space with continuous action of the Galois group $\Gamma_{E(G)}$, commuting with an action of $G(\mathbf{A}^f)$ that makes it an admissible $G(\mathbf{A}^f)$-module; in particular, the topology is irrelevant to the $G(\mathbf{A}^f)$-action.

Fix an irreducible representation π_f of $G(\mathbf{A}^f)$ defined over a finite extension $E(\pi)$ of $E(W)$.[20] For any prime λ of $E(\pi)$ (whose restriction to $E(W)$ is still denoted λ), we define

$$\begin{aligned} H^i_\lambda(S(G), \tilde{W})[\pi_f] &= Hom_{G(\mathbf{A}^f)}(\pi_f, H^i_\lambda(S(G), \tilde{W})); \\ [H_\lambda(S(G), \tilde{W})][\pi_f] &= \sum_i (-1)^i H^i_\lambda(S(G), \tilde{W})[\pi_f] \end{aligned} \tag{III.1.2}$$

the latter an object in an appropriate Grothendieck group. Given an automorphic representation π of $G(\mathbf{A})$, let ρ^i_{λ,π_f}, resp. $[\rho]_{\lambda,\pi_f}$, denote the representation (resp. virtual representation) of $\Gamma_{E(G)}$ on $H^i_\lambda(S(G), \tilde{W})[\pi_f]$ (resp. $[H_\lambda(S(G), \tilde{W})][\pi_f]$).

When $G = GL(2)_{\mathbb{Q}}$ and W is the trivial representation, $S(G)$ is the tower of elliptic modular curves, and the representations ρ^1_{λ,π_f} were originally constructed by Eichler and Shimura who identified the associated L-functions with Hecke's L-functions of modular forms of weight 2. Deligne generalized these results to arbitrary W [D68]. For general G, the determination of the representations ρ^i_{λ,π_f} and

[20] The finite parts of cohomological automorphic representations are always defined over number fields.

$[\rho]_{\lambda,\pi_f}$, as a function of π_f, is the subject of an important conjecture of Langlands [L79], extended and refined by Kottwitz [K90] in the light of Arthur's multiplicity conjectures. The conjectures were formulated together with a strategy of proof [L76, K90], based on the coincidence, first observed by Ihara [I], between the elliptic orbital integrals arising in the Selberg trace formula and the expression, provided by Grothendieck's Lefschetz formula for the number of points on Shimura varieties over a finite field of good reduction. For groups other than $GL(2)$ and its inner forms, endoscopy intervenes on both sides in the comparison between the two formulas. At a prime v of $E(G)$ of good reduction, Kottwitz obtained an expression for the characteristic polynomial of $[\rho_{\lambda,\pi_f}](Frob_v)$ on the basis of (a) natural assumptions about the parametrization of points of $S(G)$ over the residue field of v (see also [LRa]), and (b) the Fundamental Lemma for endoscopy [K90]. For $S(G)$ of PEL type, with G anisotropic modulo center and with Dynkin diagram of type A or C, Kottwitz verified the assumptions (a), thus reducing the determination of $[\rho_{\lambda,\pi_f}]$ to the Fundamental Lemma [K92b].

For a very special class of Shimura varieties attached to twisted unitary groups, Kottwitz discovered that the Fundamental Lemma for endoscopy is superfluous, and for these varieties he completely determined the characteristic polynomials of Frobenius for almost all primes of good reduction in [K92a]. This result, together with the analogue of Labesse's base change theorem II.5.4 for these groups, formed the basis for Clozel's construction of Galois representations attached to the cohomological cuspidal automorphic representations Π of $GL(n)$ over a CM field, invariant under θ and satisfying an additional local hypothesis at a finite prime [C91]. The local ℓ-adic Galois representations were subsequently completely determined at primes v of $E(G)$ prime to ℓ in [HT], with the local Langlands correspondence as a consequence. Now that the Fundamental Lemma is available, the methods described in §II.5 permit application to Shimura varieties attached to general unitary groups (anisotropic mod center) of the techniques of Kottwitz [CHL2] and of Harris-Taylor [Shin]. The results are described in the next section, along with a brief discussion of certain non-compact Shimura varieties [Mo].

Shimura varieties not of PEL type. The point counting arguments developed by Langlands and Kottwitz can be applied to Shimura varietes of PEL type, whose points over a finite field k of good reduction are identified with abelian varieties with additional structure. The isogeny classes of these abelian varieties correspond to Honda-Tate data, generalized to incorporate the PEL structure. Within each isogeny class, the isomorphism classes are parametrized by a pair of data: the first is profinite, and corresponds to the level structure away from the characteristic of k, the second is algebraic, and corresponds to the Dieudonné module of the abelian variety, again with the PEL data incorporated. At primes of bad reduction the theory of Rapoport-Zink [RZ] provides a rigid-analytic parametrization in general, that has been completely analyzed for the varieties attached to the groups analogous to GU_1 [HT,Shin], and studied in some depth by Fargues and Mantovan [FaMa] in a more general setting. When the bad reduction is of parahoric type, [RZ] defines a geometric parametrization that has been the object of extensive study; see [Rp,Hn05] for surveys of this work).

A conjecture of Langlands and Rapoport [LRa] proposes an explicit parametrization of points on general Shimura varieties, compatible with the Langlands and Kottwitz conjectures on $[\rho_{\lambda,\pi_f}]$, in terms of representations of a proalgebraic group modeled on the conjectural Tannakian category of motives over finite fields; see [Ml05] and Milne's article in [LR] for surveys of this conjecture. There has been a fair amount of work on integral models of Shimura varieties whose characteristic zero models parametrize families of abelian varieties with Hodge classes, but even in this relatively concrete case the Langlands-Rapoport conjecture seems out of reach. Practically nothing is known about the exceptional cases.

III.1.2. The case of unitary groups: construction of Galois representations.

Galois representations attached to the cuspidal automorphic representations of $GL(n)$ considered in §II.5 are realized in the cohomology of Shimura varieties attached to the unitary group U_1 constructed in Lemma II.5.5. We begin by reformulating the main spectral results of §II.5 in the language of the previous section.

Let $GU(V_1)$ denote the subgroup of $GL(V_1)$ preserving the hermitian form of II.5.5 up to a scalar (similitude) factor, necessarily rational over F. The similitude defines a homomorphism $\nu : GU(V_1) \to GL(1)_F$, and we let GU_1 be the reductive group over $\mathbb{Q}$ that is the fiber product $GL(1)_{\mathbb{Q}} \times_{GL(1)_F} GU(V_1)$ with respect to the natural inclusion $GL(1)_{\mathbb{Q}} \to GL(1)_F$ and the map ν. The Shimura variety is attached to the group GU_1 rather than to U_1.

Theorem III.1.2.1. *The hypotheses of §II.5 are in force. Let n be an even nonnegative integer.*

(a) Let Π be a cuspidal automorphic representation of $GL(n+1)_{F'}$, cohomological for $W \otimes W^{\theta}$, and satisfying $\Pi \xrightarrow{\sim} \Pi^{\theta}$. Let π_f be the unique representation of $U_1(\mathbf{A}^f)$ that transfers to Π_f and let π_f^+ denote an extension of π_f to $GU(V_1)(\mathbf{A}^f)$. If the highest weight of W is regular, then for any finite place λ of $E(\pi)$, $\rho^i_{\lambda,\pi_f^+} = 0$ unless $i = n$, in which case ρ^n_{λ,π_f^+} is an $n+1$-dimensional continuous representation of $\Gamma_{E(GU_1)}$.

(b) Let Π be a cuspidal automorphic representation of $GL(n)_{F'}$, cohomological for $W \otimes W^{\theta}$, and satisfying $\Pi \xrightarrow{\sim} \Pi^{\theta}$. Choose χ as in II.5.6, let $\pi_f(\Pi_f, \chi_f)$ be the representation of $U_1(\mathbf{A}^f)$ defined there, and let π_f^+ denote an extension of π_f to $GU(V_1)(\mathbf{A}^f)$. If the highest weight of W is regular, or if U_∞ is a product of compact unitary groups, then for any finite place λ of $E(\pi)$, $\rho^i_{\lambda,\pi_f^+} = 0$ unless $i = n$, in which case ρ^n_{λ,π_f^+} is an n-dimensional continuous representation of $\Gamma_{E(GU_1)}$.

Parts (a) and (b) follow with minimal effort from Theorems II.5.4 and II.5.6, respectively.[21] Henceforward we assume $F' = \mathcal{K} \cdot F$ for some imaginary quadratic field $\mathcal{K}$. The datum defining $S(GU_1)$ depends on the choice of a CM type Σ for F', which we have not yet specified; henceforward we take Σ to be the set of extensions to F' of a fixed complex embedding of $\mathcal{K}$. For reasons we will not explain, this guarantees that $E(GU_1) \xrightarrow{\sim} F'$, so the Galois representations of III.1.2.1 are representations of $\Gamma_{F'}$.

We fix a level subgroup $K_f \subset GU_1(\mathbf{A}^f)$. The Shimura variety $S(GU_1)_{K_f}$ is of PEL type and the methods of [K92b] apply at places of good reduction. At places of bad reduction, a synthesis of the constructions of [K92b] and [HT], due to Shin,

[21] The article [Shin] uses a method of Taylor [HT] to construct Galois representations of dimensions $n+1$ and n in the above situations, without the regularity hypothesis on W nor the hypothesis II.5.3 (1) and (2), and without the precise multiplicity assertions of [CHL1].

permits the determination of the representation of the local Galois group Γ_v on $[\rho_{\lambda,\pi_f^+}]$ at any place v of F' prime to λ. The following theorem summarizes the main results of [CHL2], [Shin], and [CH].

Theorem III.1.2.2. *Let n be a positive integer and let Π be a cohomological cuspidal automorphic representation of $GL(n)_{F'}$ satisfying $\Pi \xrightarrow{\sim} \Pi^\theta$. There is a compatible family of continuous representations*

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

as λ runs through non-archimedean completions of a certain number field $E(\Pi)$, satisfying

(a) *$\rho_{\Pi,\lambda}$ is* **geometric** *in the sense of Fontaine and Mazur [FM]. More precisely, $\rho_{\Pi,\lambda}$ is unramified outside a finite set of places of F', namely the set of places of ramification of Π and the set of primes of F' dividing the residue characteristic ℓ of λ. Moreover, at places dividing ℓ, $\rho_{\Pi,\lambda}$ is de Rham, in Fontaine's sense, and in particular is Hodge-Tate.*
(b) *$\rho_{\Pi,\lambda}$ is* **HT regular***: the Hodge-Tate weights of $\rho_{\Pi,\lambda}$ have multiplicity one.*
(c) *There is a non-degenerate bilinear pairing*

$$\rho \otimes \rho \ \to E(\Pi)_\lambda \otimes \mathbb{Q}_\ell(1-n).$$

This correspondence has the following properties:

(i) *For any finite place v prime to the residue characteristic ℓ of λ,*

$$(\rho_{\Pi,\lambda} \mid_{\Gamma_v})^{ss} = \mathcal{L}(\Pi_v)^{ss}.$$

Here Γ_v is a decomposition group at v and $\mathcal{L}$ is the **normalized** *local Langlands correspondence (denoted $r_\ell(\Pi_v)$ in [HT]). The superscript ^{ss} denotes semisimplification. If the infinitesimal character of Π_∞ is sufficiently regular, we even have*

$$(\rho_{\Pi,\lambda} \mid_{\Gamma_v})^{Frob-ss} = \mathcal{L}(\Pi_v),$$

where the superscript on the left denotes Frobenius semisimplification.
(ii) *The representation $\rho_{\Pi,\lambda} \mid_{\Gamma_v}$ is de Rham for any v dividing ℓ and the Hodge-Tate numbers at v are explicitly determined by the archimedean weight of Π_∞, or equivalently by the coefficients W, by the formula given in* [HT,Theorem VII.1.9].
(iii) *If Π_v is unramified then $\rho_{\Pi,\lambda} \mid_{\Gamma_v}$ is crystalline. If Π_v has an Iwahori-fixed vector then $\rho_{\Pi,\lambda} \mid_{\Gamma_v}$ is semi-stable (in Fontaine's sense).*

The statement is taken from [CH,Theorem 3.2.5]. There is an analogous result when Π is a cohomological cuspidal automorphic representation of $GL(n)_F$ that is *self-dual* (i.e., its base change to $GL(n)_{F'}$ is θ-invariant), cf. [CH, Theorem 4.2].

Open questions.

(i) (Generalized Ramanujan Conjecture) It was proved in [HT] that, for the Π treated there, the local representation Π_v is tempered for all v. This was verified more generally in [Shin] for representations satisfying a weak regularity condition. For the remaining representations, the methods of [CH] apparently imply that Π_v is tempered for almost all v, but something more is needed at the missing places.

(ii) (Irreducibility) It is expected that $\rho_{\Pi,\lambda}$ be *irreducible* for all λ. This is only known in general for $n \leq 3$ [BR92] and in some cases for $n = 4$. The most direct route to the general case apparently requires a considerable strengthening of the results of §III.2.

Remark on complex multiplication. The case $n+1 = 1$, corresponding to $\dim V_1 = 1$, is precisely the main theorem of complex multiplication due to Shimura and Taniyama [ST]. One should bear in mind, however, that the theory of complex multiplication is required to define the canonical models on general $S(GU(V_1))$. The result is not circular, however; the theory of complex multiplication can be reproved in the setting of 0-dimensional Shimura varieties of PEL type, using their realization as moduli spaces and applying the methods of [K92b] or [HT].

Remark on the proofs. First assume F' contains an imaginary quadratic field $\mathcal{K}$ as above, so that the spaces constructed of Theorem III.1.2.1 (a) carry representations of $\Gamma_{F'}$. For n odd, we apply III.1.2.1(a). The representation $\rho_{\Pi,\lambda}$ is obtained as a certain abelian twist of the dual of the representation ρ^n_{λ,π_f^+} constructed in the cohomology of $S(GU_1)$; cf. [CHL2, Lemma 4.4.7] for the precise formula. The proof at almost all primes at which Π is unramified is given in [CHL2], following Kottwitz and applying the Fundamental Lemma where it was assumed in [K90], provided the infinitesimal character is sufficiently regular. A proof valid at all places prime to the residue characteristic ℓ of λ, without the regularity condition, is due to Shin [Shin].

For n even, we apply III.1.2.1(b), generalizing an idea due to Blasius and Rogawski [BR]. Again, [CHL2] follows Kottwitz and works only at unramified primes and for sufficiently regular infinitesimal character at all primes of F' dividing ∞; [Shin] gives an independent proof valid at all v prime to ℓ and assumes only a weak regularity condition on Π_v at a single archimedean place v.

It remains to treat the case of n even, when Π is cohomological for the representation $W \otimes W^\theta$ when the highest weight of W does not satisfy a special regularity condition. The case of W trivial, in particular, is not covered by the results of [CHL2] and [Shin]. Indeed, in this case $\rho_{\Pi,\lambda}$ *cannot* in general be realized in the cohomology of a Shimura variety, so the trace formula arguments do not apply. To complete the proof of the theorem, one needs to apply congruence arguments. This phenomenon already arose when $n = 2$ when $[F : \mathbb{Q}]$ is even [BR]. The only known constructions of the Galois representations attached to Hilbert modular forms of weight $(2, 2, \ldots, 2)$ are due to Wiles [W88], when the form is ordinary at λ, and in general to Taylor [T88]. The method of [CH] generalizes that of Wiles; the Hida families used in [W88] are replaced by the families of Galois (pseudo)-representations over the *eigenvarieties* constructed by Chenevier in [Che], whose relevant properties were established in [BC]. Roughly speaking, the eigenvariety parametrizes automorphic representations of the totally definite unitary group U_0. Labesse's theorem II.5.4 asserts that Π descends to a (unique) automorphic representation of U_0 that corresponds to (one or more) points on the eigenvariety. The idea is to approximate one of these points z λ-adically by a sequence of points z_α corresponding to Π that do satisfy the regularity condition; then the corresponding λ-adic representations ρ_{z_α} converge to the desired representation $\rho_{\Pi,\lambda}$. The de Rham condition, and the stronger ℓ-adic ramification conditions indicated in (iii), are proved using the full strength of Fontaine's theory; the proof is based on the generalization in [BC] of Kisin's theory of analytic continuation of crystalline periods. These conditions are essential in the applications to deformation theory.

Two obstacles remain. In the first place, we need to remove the hypothesis that F' contains an imaginary quadratic field. In the second place, it may be that Π_v is highly ramified (not contained in a principal series) for some v dividing ℓ; in that case Π does not define a point on the eigenvariety. Both obstacles can be removed by a solvable base change and patching argument that originally goes back to Blasius and Ramakrishnan; the form used here it is due to Sorensen [So]. This completes the sketch of the proof.

Theorem III.1.2.2 is the generalization of the theorems of Eichler, Shimura, and Deligne to θ-invariant cohomological automorphic representations of $GL(n)$ over any CM field. It is a direct descendent of Takagi's existence theorem: the cuspidal automorphic representation Π is the non-abelian version of the congruence conditions of class field theory, and the kernels of the reductions mod λ^n of the homomorphisms $\rho_{\Pi,\lambda}$ are finite non-abelian extensions of F' that generalize the class fields whose existence was proved by Takagi.

Note that condition (b) of III.1.2.2 is a severe restriction on the kind of representations that arise in this way; in particular the $\rho_{\Pi,\lambda}$ are very far from the Artin representations $\mathcal{G}_n(F')_{fin}$ mentioned in the introduction. Moreover, the restriction to CM and totally real fields is totally unnatural from the standpoint of generalizing class field theory, and indeed Langlands' functoriality conjectures involve no such restriction.

Nevertheless, it seems fair to say that Theorem III.1.2.2 is very close to the outer limit of what can be achieved by studying the cohomology of Shimura varieties. Langlands and Kottwitz have conjectured explicit formulas for the ρ_{λ,π_f} discussed in the previous section. At least for PEL types it is likely that the strategy outlined in [K90, K92b], together with the methods of Morel's recent work on non-compact Shimura varieties [Mo], should prove the Langlands and Kottwitz conjectures completely at unramified primes. Indeed, for groups of type A and C, most of the work has already been done by Kottwitz, as we have seen in [CHL2]. Since the groups attached to PEL Shimura varieties are classical groups, combining Arthur's results on functorial transfer with Shin's results on bad reduction of the Shimura varieties $S(GU_1)$ should prove most of what is expected at ramified primes as well. But the list of Galois representations expected to arise in the cohomology of Shimura varieties is nearly exhausted by Theorem III.1.2.2. It should be possible to obtain additional Galois representations, including some that are geometric in the sense of Fontaine-Mazur, by congruence arguments. But at present no one knows where to begin to develop a general theory with the scope of abelian class field theory (see however §III.4, below). This is to be contrasted with the very different situation of function fields – the base field F or F' is replaced by the field of rational functions on an algebraic curve over a finite field – where Lafforgue, extending the work of Drinfel'd on $GL(2)$, completely proved the Langlands correspondence for $GL(n)$ [Lf].

On the other hand, remaining within the framework of the cohomology of Shimura varieties, one can ask whether every representation that looks like a $\rho_{\Pi,\lambda}$ is in fact associated to an automorphic representations. This is the problem of *reciprocity*, the starting point for most of the arithmetic applications of automorphic forms and the subject of the next section.

III.2. Automorphic lifting theorems

Let F be a number field and G a connected reductive group over F. The Langlands functoriality conjectures are sufficiently general to account for all homomorphisms of the Galois group Γ_F to the complex L-group of G in terms of automorphic representations of G, and thus include a version of non-abelian class field theory for finite extensions of F. Very little has been proved in this direction, and most of that is only for the group $GL(2)$. On the other hand, we have seen that many automorphic representations of groups attached to Shimura varieties, and of $GL(n)$ when F is totally real or a CM field, give rise to compatible families of finite-dimensional λ-adic representations of Γ_F.

Not every finite-dimensional λ-adic representation is expected to be attached to automorphic forms. For example, the representations constructed by Ramakrishna [Rk] and Khare, Larsen, and Ramakrishna [KLR] are ramified at infinitely many primes, whereas an ℓ-adic Galois representation attached to an automorphic representation Π can at most be ramified at the finite set of places where Π is ramified, together with primes dividing ℓ. On the other hand, the theory of p-adic modular forms provides continuous p-adic analytic families of p-adic Galois representations with a given determinant. These representations are ramified at a finite set of primes, but there are uncountably many of them, whereas the set of cuspidal automorphic representations with fixed central character (which corresponds to the determinant of the Galois representation under Langlands reciprocity) is countable. Both pathologies are already known, with appropriate modifications, for 1-dimensional ℓ-adic representations.

Theorem III.1.2.2 asserts that the λ-adic representations attached to the Π considered in §III.1 satisfy the *geometric* condition of Fontaine and Mazur. This condition was introduced in [FM] in connection with the authors' conjectures on representations of this type, including what has come to be known as *the Fontaine-Mazur Conjecture*, that 2-dimensional geometric ℓ-adic representations of $\Gamma_{\mathbb{Q}}$ are attached to modular forms. Geometricity in this sense seems to be the appropriate condition for ensuring that a general λ-adic representation is attached to automorphic representations. To my knowledge, a full conjecture to this effect has not been proposed in print (see [T02a], for example), and I will not do so here. The remainder of §III will be concerned instead with the following more limited conjecture.

Generalized Fontaine-Mazur conjecture. *Let F' be a CM field, n a positive integer, and L be an ℓ-adic field. Let*

$$\rho : \Gamma_{F'} = Gal(\overline{\mathbb{Q}}/F') \to GL(n, L)$$

be a continuous absolutely irreducible representation that satisfies (a)-(c) of Theorem III.1.2.2. Then there is a cuspidal automorphic representation Π of $GL(n)_{F'}$, satisfying the hypotheses of III.1.2.2, a place λ of $E(\Pi)$, and an inclusion $\iota : E(\Pi)_\lambda \hookrightarrow L$, such that ρ is equivalent to $\rho_{\Pi,\lambda} \otimes_{E(\Pi)_\lambda,\iota} L$.

The conjecture has been proved in almost all cases[22] in its original context of 2-dimensional representations of $\Gamma_{\mathbb{Q}}$. Partial results have been obtained recently

[22] The main reference is [KiFM], in which ℓ is assumed odd and several additional hypotheses

for general CM fields. These results are all based on Wiles' approach to arithmetic deformation rings, to which we now turn.

III.2.1. Deformations of Galois representations.

Let F' be a CM field as above with totally real subfield F. Let $\mathcal{O}$ be the ring of integers in a finite extension of $\mathbb{Q}_\ell$, with residue field k and maximal ideal $\mathfrak{m}$. Let $\rho : \Gamma_{F'} \to GL(n,\mathcal{O})$ be a continuous representation and let $\bar{\rho} : \Gamma_{F'} \to GL(n,k)$ be its reduction modulo $\mathfrak{m}$. Fix a finite set of primes S of F including the set S_ℓ of primes dividing ℓ and all primes at which $\bar{\rho}$ is ramified. Let $\Gamma_{F',S}$ denote the Galois group of the maximal extension of F' unramified outside S, and define $\Gamma_{F,S}$ similarly.

Definition III.2.1.1. *Let A be a noetherian local $\mathcal{O}$-algebra with maximal ideal $\mathfrak{m}_A$ and residue field k. A* **lifting** *of $\bar{\rho}$ to A (understood to be unramified outside S) is a homomorphism $\tilde{\rho} : \Gamma_{F',S} \to GL(n,A)$ together with an isomorphism*

$$\tilde{\rho} \pmod{\mathfrak{m}_A} \xrightarrow{\sim} \bar{\rho}$$

compatible with the natural map $\mathcal{O}/\mathfrak{m} = A/\mathfrak{m}_A = k$. A **deformation** *of $\bar{\rho}$ to A is an equivalence class of liftings, where two liftings are equivalent if one can be obtained from the other by conjugation by an element of $GL(n,A)$ congruent to the identity modulo $\mathfrak{m}$.*

Let $Def_{\bar{\rho},S}$ be the functor on the category of Artinian local $\mathcal{O}$-algebras that to A associates the set of equivalence classes of deformations of $\bar{\rho}$ to A (unramified outside S). This can be extended to a (pro)-functor on the category of noetherian local $\mathcal{O}$-algebras, and the starting point of the theory is Mazur's theorem. Let $ad(\bar{\rho})$ denote the natural representation of $\Gamma_{F',S}$ on $End(\bar{\rho})$.

Theorem III.2.1.2 [Mz89]. *Suppose $\bar{\rho}$ is absolutely irreducible (or more generally $End_{\Gamma_{F',S}}(\bar{\rho} \otimes \bar{k}) = \bar{k}$). Then $Def_{\bar{\rho},S}$ is (pro)representable by a noetherian local $\mathcal{O}$-algebra $R_{\bar{\rho},S}$, with maximal ideal $\mathfrak{m}_R$. Moreover, the Zariski tangent space $Hom(\mathfrak{m}_R/(\mathfrak{m}_R)^2, k)$ of this functor is naturally isomorphic to the Galois cohomology group $H^1(\Gamma_{F',S}, ad(\bar{\rho}))$.*

Mazur's theory has been studied by many people and has been the subject of numerous survey articles, including several articles in [CSS]; there is also a very brief account in [H3]. Wiles' paper [Wi] on Fermat's Last Theorem created a systematic approach to proving statements like the Fontaine-Mazur Conjecture, in three steps. Let ρ be a representation as in the statement of the conjecture, and let $\mathcal{O}$ be the ring of integers of L. Since $\Gamma_{F'}$ is compact, its image under ρ fixes an $\mathcal{O}$-lattice, and thus defines a residual representation $\bar{\rho}$.

1. Step one is to prove that $\bar{\rho}$ admits a lifting $\tilde{\rho}$ to an *automorphic* Galois representation, i.e., $\tilde{\rho}$ is isomorphic to $\rho_{\Pi,\lambda}$ for some cuspidal Π.
2. Step two is to recognize that the obstruction to lifting $\rho \pmod{\mathfrak{m}^2}$ to an automorphic $\tilde{\rho}$ can be interpreted in terms of the Zariski tangent space; in other words the obstruction can be interpreted in terms of Galois cohomology.

---

are assumed – see the statement of the Theorem on p. 642. I have been advised not to worry about condition (2). The article [Ki2] extends the results of [KiFM] to 2-adic representations but only considers potentially Barsotti-Tate representations.

(3) At this point one is faced with an infinite sequence of cohomological obstructions to automorphic lifting modulo increasing powers of $\mathfrak{m}$. Step three is to recognize that, thanks to powerful techniques of commutative algebra, one need not worry about higher obstructions.

This is not how Wiles' method is conventionally presented. In its current standard form, which originated in the Taylor-Wiles paper [TW], step 2 is replaced by a counting argument based on Galois cohomological duality, and the burden of step 3 is mainly shifted to the study of modules of automorphic forms over integral Hecke algebras. But it may be helpful to focus on the underlying topological argument, especially since the Taylor-Wiles method is by now quite familiar (cf. the articles just cited, for example).

In what follows I will concentrate on relating the Taylor-Wiles method in dimensions > 2 to the automorphic theory presented in §II. The key notion that links the two theories is that of *Hecke algebra*, especially the unramified Hecke algebras introduced in §I.2.4.

The functor of (III.2.1.2) allows arbitrary ramification at primes in S, and is generally too coarse for applications. A collection of restrictions on ramification at S, which we denote $\mathcal{S}$, define a *deformation problem* if they can be expressed in terms of the localization maps

$$H^1(\Gamma_{F',S}, ad(\bar{\rho})) \to H^1(\Gamma_v, ad(\bar{\rho})),\ v \in S$$

where Γ_v is the local decomposition group, and if the functor they define remains representable. At primes in S_ℓ, in particular, one only wants to consider $\tilde{\rho}$ that are de Rham. For $n > 2$ one only knows how to interpret this condition cohomologically if $\tilde{\rho}$ is in fact crystalline, and if ℓ is large relative to the Hodge-Tate numbers. One also restricts attention to $\tilde{\rho}$ satisfying the polarization condition (c) of (III.1.2.2); this is both representable and admits a cohomological interpretation, in that it allows an extension of $ad(\bar{\rho})$ from $\Gamma_{F',S}$ to $\Gamma_{F,S}$. The Zariski tangent space to the resulting deformation problem, and thus to the deformation ring $R_{\bar{\rho},\mathcal{S}}$ (note that S has been replaced by $\mathcal{S}$), is a *Selmer group* $H^1_{\mathcal{S}}(\Gamma_{F,S}, ad(\bar{\rho}))$ contained naturally in $H^1(\Gamma_{F,S}, ad(\bar{\rho}))$. See [H3, §3], for more details.[23]

For Mazur's theorem it is essential that $End_{\Gamma_{F,S}}(\bar{\rho} \otimes \bar{k}) = \bar{k}$. In the applications in higher dimensions, one only knows how to treat $\bar{\rho}$ that are absolutely irreducible. Reducible 2-dimensional representations of $\Gamma_{\mathbb{Q}}$ have been studied in the paper of Skinner-Wiles [SW], which depends on methods that for the moment are not available in higher dimensions.

III.2.2. Hecke algebras. We now start with a cuspidal automorphic representation Π_0 of $GL(n, F')$, satisfying the hypotheses of §III.1.2. Recall the group U_0 of (II.5.5). Theorem II.5.4 asserts that Π_0 corresponds to a unique automorphic representation π_0 of U_0 under the simplifying hypotheses II.5.3, which we will assume for simplicity. For purposes of exposition we take coefficients W to be the trivial representation until the end of this section; this allows us to define an ℓ-integral version $\mathbb{Z}_\ell$ without additional choices.

Assume $\pi_0^{K_f} \neq 0$ for some level subgroup $K_f = \prod K_v \subset U_0(\mathbf{A}^f)$. Then π_0 contributes to the cohomology $H^0({}_{K_f}S(U_0), \mathbb{Z}_\ell)$ (no˜necessary on the coefficients), a free $\mathbb{Z}_\ell$-module of finite rank. Let S be the finite set of v at which K_v is not a

[23] The reader seeking all the details will have to look at [CHT], [T08], [Ge], and [Gu].

hyperspecial maximal compact subgroup, and let $\mathcal{H}_\ell^S$ denote the restricted tensor product

$$\bigotimes_{v \notin S}{}' \mathcal{H}_{\mathbb{Z}}(U_0(F_v), K_v) \otimes \mathbb{Z}_\ell$$

where $\mathcal{H}_{\mathbb{Z}}$ is defined as in I.2.4.3. The algebra $\mathcal{H}_\ell^S$ acts by right convolution on $H^0({}_{K_f}S(U_0), \mathbb{Z}_\ell)$ and we let $\mathbb{T}_{K_f} \subset End(H^0(S(U_0)_{K_f}, \mathbb{Z}_\ell))$ denote the subalgebra generated by $\mathcal{H}_\ell^S$, a finite commutative reduced $\mathbb{Z}_\ell$ algebra.[24]

Fix a prime λ of the coefficient field $E(\Pi_0)$ and let $\rho_0 = \rho_{\Pi_0,\lambda}$, with coefficients in the ℓ-adic integer ring $\mathcal{O}$, which we allow to grow in what follows. The natural map

$$H^0(S(U_0)_{K_f}, \mathbb{Z}_\ell) \to \pi_{0,f}^{K_f} \otimes Hom_{\mathcal{H}_\ell^S}(\pi_{0,f}^{K_f}, H^0_\lambda(S(U_0), E(\Pi_0)_\lambda)) \quad \text{(III.2.2.1)}$$

defines a prime ideal $\mathfrak{P}_0 \subset \mathbb{T}_{K_f}$, the kernel of the character $\chi_0 : \mathbb{T}_{K_f} \to E(\Pi_0)_\lambda$ by which the Hecke algebra acts on the right hand side of (III.2.2.1). Since the image of χ_0 is contained in a field, $\mathfrak{P}_0$ is contained in a unique maximal ideal $\mathfrak{m}_0$. On the other hand, it follows directly from Theorem III.1.2.2 (i) that

(III.2.2.2) Unramified local reciprocity. For any $v \notin S$, the characteristic polynomial P_{v,ρ_0} of $\rho_0(Frob_v)$ has coefficients in $\chi_0(\mathbb{T}_{K_f})$.

The coefficients of P_{v,ρ_0} can be made explicit in terms of the Satake transform (cf. [CHT,Cor. 3.1.2]). In this version, III.2.2.2 is the exact generalization of the Eichler-Shimura congruence relation. For our purposes III.2.2.2 suffices to establish

Lemma III.2.2.3. *Let Π be a cuspidal automorphic representation of $GL(n, F')$ satisfying the hypotheses of §III.1.2, with coefficients W trivial. Let π be the corresponding representation of U_0 and assume $\pi^{K_f} \neq 0$; let χ_π be the corresponding character of $\mathbb{T}_{K_f}$. Then* $\ker(\chi_\pi) \subset \mathfrak{m}_0$ *if and only if the semisimplifications of $\bar\rho_0$ and $\bar\rho_{\Pi,\lambda}$ are isomorphic. In particular, if $\bar\rho_0$ is absolutely irreducible,* $\ker(\chi_\pi) \subset \mathfrak{m}_0$ *if and only if $\rho_{\Pi,\lambda}$ is a lifting of $\bar\rho_0$, and hence defines a (classifying) map*

$$R_{\bar\rho_0,\mathcal{S}} \to \mathcal{O}$$

where $\mathcal{O}$ has been expanded if necessary to contain the integer ring of $E(\Pi)_\lambda$ and where $\mathcal{S}$ defines an appropriate deformation problem.[25]

Assume henceforward that $\bar\rho_0$ *is absolutely irreducible.* Define $\mathbb{T}_{\rho_0,\mathcal{S}}$ to be the completion of $\mathbb{T}_{K_f}$ at the maximal ideal $\mathfrak{m}_0$. Applying Lemma III.2.2.3 and a theorem due to Carayol [Ca] to the set of all Π such that $\ker(\chi_\pi) \subset \mathfrak{m}_0$, one obtains a classifying map

$$R_{\bar\rho_0,\mathcal{S}} \to \mathbb{T}_{\rho_0,\mathcal{S}} \quad \text{(III.2.2.4)}$$

of noetherian local $\mathcal{O}$-algebras and thus a map

$$\mathfrak{m}_R/(\mathfrak{m}_R)^2 \to \mathfrak{m}_{\mathbb{T}}/(\mathfrak{m}_{\mathbb{T}})^2 \quad \text{(III.2.2.5)}$$

[24]In [CHT] and subsequent articles, one only takes the tensor product over v split in F'/F, which is adequate for the applications but does not generalize to other groups. I haven't checked that the present formulation has all the properties claimed here, although I have no doubt that the claims are true.

[25]This last step is an oversimplification; explaining how one makes the appropriate choices for $\mathcal{S}$ accounts for much of the length of [CHT], for example.

of Zariski cotangent spaces. It follows from III.2.2.2 that this map is surjective.

On the other hand, any deformation ρ of $\bar{\rho}_0$ of type $\mathcal{S}$, with coefficients in $\mathcal{O}$, determines a classifying map

$$c_\rho : R_{\bar{\rho}_0,\mathcal{S}} \to \mathcal{O}.$$

The deformation ρ is automorphic if and only if c_ρ factors through (III.2.2.4). More precisely, if c_ρ factors through (III.2.2.4) if and only if ρ corresponds to an automorphic representation of U_0. The relation to automorphic representations of $GL(n)$ is guaranteed by Labesse's Theorem II.5.4. The passage to $GL(n)$ involves some subtleties; we return to this point below. The kernel of (III.2.2.5) measures the first-order obstruction to (III.2.2.4) being an isomorphism. We can now explain Steps 2 and 3 of Wiles' approach to deformations, as sketched above: to prove that all deformations of $\bar{\rho}_0$ are automorphic it suffices with the vanishing of this obstruction, interpreted in terms of an appropriate Selmer group, as well as all higher obstructions. The next section describes some recent results in this direction. We conclude this section by observing that all of the statements in this section remain valid for arbitrary coefficients W.

III.2.3. Results on automorphic lifting.

The deformation theory of 2-dimensional Galois representations, initiated by Wiles [Wi95] and pursued in the articles [TW], [D], [F], [BCDT] and [KiMF], among others, has culminated in the (nearly) complete solution of the original Fontaine-Mazur conjecture by Kisin [KiFM], and a second solution by Emerton (in preparation), both based on the p-adic Langlands correspondence (see §IV.3), and in the proof of Serre's Conjecture on 2-dimensional *modular* representations of $\Gamma_{\mathbb{Q}}$ by Khare and Wintenberger, together with a great deal of the 2-dimensional Artin Conjecture over $\mathbb{Q}$. These developments will undoubtedly be treated thoroughly in Khare's talk and I will concentrate on what is known in arbitrary dimension.

The general framework sketched in the previous section was developed at length in [CHT] in the setting of the Galois representations of [C91,K92a,HT]. These papers were written before the proof of the Fundamental Lemma and therefore relied on the twisted unitary groups studied in [K92a]. In [CHT] the only deformations considered were those that are crystalline at primes above ℓ and *minimal* at other primes; roughly; at $v \nmid \ell$, it is assumed that $\tilde{\rho}$ is no more ramified than $\bar{\rho}$. The article [T08] completed [CHT] by eliminating the minimality restriction, completely settling the problems of level-raising and level-lowering as they apply to proving that Galois representations are automorphic. The more recent papers of Guerberoff and Geraghty [Gu,Ge] extended this framework in various ways, in particular taking advantage of Labesse's version of base change in Theorem II.5.4. Since the conclusions are identical in all of these papers, and the hypotheses differ only slightly, they will be presented here as variants of a single theorem.

Modularity Lifting Theorems. *Let F' be a CM field and let Π_0 be a cuspidal automorphic representation of $GL(n)_{F'}$ satisfying $\Pi_0 \xrightarrow{\sim} \Pi_0^\theta$, and assume Π_0 is cohomological with respect to $W \otimes W^\theta$ for some irreducible representation W of $R_{F/\mathbb{Q}}U_0$. Let $\rho_0 = \rho_{\Pi_0,E(\Pi_0)_\lambda}$ be as in §III.2.2, where λ is a prime of $E(\Pi_0)$ of residue characteristic $\ell > n$. We assume the coefficients $\mathcal{O}$ of ρ_0 are sufficiently big in what follows, and let k denote the residue field of $\mathcal{O}$.*

Let $\rho : \Gamma_{F'} \to GL(n, \mathcal{O})$ be a deformation of $\bar{\rho}_0$ that satisfies conditions (a), (b), and (c) of Theorem III.1.2.2. Suppose moreover that the Hodge-Tate weights of ρ

are the same as those of ρ_0. Assume every prime v of F dividing ℓ splits in F'. Finally assume
(IM) *$\bar{\rho}_0$ is absolutely irreducible, the image of $\bar{\rho}_0$ in $GL(n,k)$ is "big" in the sense of [CHT, Def. 2.5.1], and the fixed field under the kernel of the representation $ad(\bar{\rho}_0)$ does not contain the ℓth roots of unity.*

III.2.3.1. [CHT, Theorem 4.4.2] *Assume*

(i) *ℓ is unramified in F and $\Pi_{0,v}$ is unramified at all primes of F' dividing ℓ.*
(ii) *ℓ is big relative to the highest weight of W (cf. condition 4 of [CHT, 4.4.2]) and ρ is crystalline at each prime v dividing ℓ.*
(iii) *There is a prime v' of F' such that $\Pi_{0,v'}$ is a discrete series representation of $GL(n, F'_{v'})$.*
(iv) *The deformation condition $\mathcal{S}$ is minimal at primes in S not dividing ℓ (cf. conditions 5, 6 of [CHT, 4.4.2]).*

Then ρ is automorphic. Indeed, the map (III.2.2.4), where $\mathcal{S}$ corresponds to deformations crystalline at ℓ and minimally ramified elsewhere, is an isomorphism.

III.2.3.2. [T08, Theorem 5.2] *Assume only conditions (i)-(iii) of III.2.3.1. Then ρ is automorphic. Indeed, the map deduced from (III.2.2.4)*

$$R^{red}_{\bar{\rho}_0,\mathcal{S}} \to \mathbb{T}_{\rho_0,\mathcal{S}} \tag{*}$$

where the superscript ^{red} denotes the quotient by the ideal of nilpotent elements and $\mathcal{S}$ corresponds to deformations crystalline at ℓ, is an isomorphism.

III.2.3.3. [Gu, Theorem 5.1] *Assume only conditions (i)-(ii) of III.2.3.1. Then ρ is automorphic. Indeed, the map (*) of III.2.3.2 is an isomorphism.*

III.2.3.4. [Ge, Theorem 5.3.2] *Assume ρ is ordinary[26] at all primes of F' dividing ℓ. Then ρ is automorphic. Indeed, the map (*) of III.2.3.2 is an isomorphism.*

Theorems III.2.3.3 and III.2.3.4, each of which comes with an analogue adapted to self-dual cohomological automorphic representations of totally real fields, represent the current state of knowledge regarding steps 2 and 3 of Wiles' program in dimensions > 2. The remaining restrictions are of two types. The conditions on the image of $\bar{\rho}_0$ are indispensable to the application of the Taylor-Wiles method, which applies Chebotarev density to control the size of the global Galois cohomology with coefficients in $ad(\bar{\rho}_0)$. The "big" condition, in particular, seems to be generic: Snowden and Wiles have recently proved [SnW] that "in a sufficiently irreducible compatible system" of ℓ-adic representations, "the residual images are big at a density one set of primes." The second set of restrictions concerns ramification at ℓ. For the moment, one can only treat crystalline ρ whose Hodge-Tate weights are small relative to ℓ, or ordinary ρ.

III.3. The method of potential automorphy

The present section considers step 1 of Wiles' strategy. While no techniques are currently available for carrying this out in practice except for very special $\bar{\rho}$, Taylor discovered an extremely powerful approach to the question if one is willing to settle for proving that a given ρ is *potentially automorphic*, that is, it becomes

[26] This corresponds to what Hida calls *nearly ordinary*.

automorphic over a finite extension of the original base field. In [T02, T06] he used this approach to show that the 2-dimensional Galois representations appearing in the Fontaine-Mazur conjecture are potentially automorphic. Using a moduli space studied in detail by Dwork, then rediscovered by physicists [COGP] in connection with mirror symmetry, [HST] proved the first potential automorphy theorems in higher dimension. These theorems have been strengthened and extended in [GHK], [BL1]. The strongest forms of the theorems will be presented below, along with some of their applications.

Rather surprisingly, Taylor's potential automorphy theorem provides the first step in the Khare-Wintenberger proof of Serre's conjecture, and thus leads via Kisin's work to the proof of the full Fontaine-Mazur conjecture for $GL(2)_{\mathbb{Q}}$. At present this line of reasoning has not been extended to other base fields, nor to representations of higher dimension.

III.3.1. Moduli spaces and potential automorphy.

Consider the equation

$$(f_t) \qquad f_t(X_1,\ldots,X_N) = (X_1^N + \cdots + X_N^N) - NtX_1\ldots X_N = 0,$$

where t is a free parameter. This equation defines an $N-2$-dimensional hypersurface $Y_t \in \mathbb{P}^{N-1}$ and, as t varies, a family:

$$\begin{array}{ccc} Y & \subset & \mathbb{P}^{N-1}\times\mathbb{P}^1 \\ & \searrow & \downarrow \\ & & \mathbb{P}^1_t \end{array}$$

Because f_t is of degree N, it follows from standard calculations that, provided it is non-singular, Y_t is a *Calabi-Yau variety*: its canonical bundle is trivial. When $N = 5$, Y is a family of quintic threefolds in $\mathbb{P}^4$, the case given special attention in [COGP]. When $t = \infty$ Y_t is the union of coordinate hyperplanes, whereas when $t = 0$, Y_0 is the Fermat hypersurface

$$X_1^N + \cdots + X_N^N = 0,$$

surprisingly useful in the applications. In any characteristic prime to N, an elementary calculation shows the *Dwork family* $\pi : Y \to \mathbb{P}^1$ is smooth over $T_0 = \mathbb{P}^1 \setminus \{\infty, \mu_N\}$.

Let μ_N denote the group of N'st roots of unity,

$$H = \mu_N^N/\Delta(\mu_N),$$

where Δ is the diagonal map, and let

$$H_0 = \{(\zeta_1,\ldots,\zeta_N) \mid \prod_i \zeta_i = 1\}/\Delta(\mu_N) \subset H.$$

The group H_0 acts on each Y_t and defines an action on the fibration $Y/\mathbb{P}^1$, and thus on the local system $R^{N-2}\pi_*\mathbb{Z}/M\mathbb{Z}$ on T_0 for any positive integer M. Assume henceforward M is prime to N. Then $R^{N-2}\pi_*\mathbb{Z}/M\mathbb{Z}[\mu_N]$ decomposes as the direct sum of H_0-isotypic components, indexed by the set $\widehat{H_0}$ of characters of H_0. The elements of $\widehat{H_0}$ can be identified with

(III.3.1.1) $$\underline{a} = \{(a_0, \ldots, a_n) \in (\mathbb{Z}/N\mathbb{Z}) \mid \sum a_i = 0\}/\Delta(\mathbb{Z}/N\mathbb{Z}),$$

and we write

$$R^{N-2}\pi_*\mathbb{Z}/M\mathbb{Z}[\mu_N] = \oplus_{\underline{a}} V^{\underline{a}}[M].$$

Poincaré duality on $R^{N-2}\pi_*\mathbb{Z}/M\mathbb{Z}$ induces a perfect pairing $V^{\underline{a}}[M] \otimes V^{-\underline{a}}[M] \rightarrow \mathbb{Z}/M\mathbb{Z}(2-N)$. In particular, when $\underline{a} = 0$, the H_0-invariants in $R^{N-2}\pi_*\mathbb{Z}/M\mathbb{Z}$ carry a perfect pairing which is alternating if N is odd and symmetric if N is even. More generally, if $\underline{a}$ differs from $-\underline{a}$ by a permutation of the indices, Katz shows in [KaDw,10.3] that $V^{\underline{a}}[M]$ has a perfect pairing $<,>_{\underline{a}}$ with parity $(-1)^N$. Call this the *dual case*.

If $\lambda \nmid N$ is a prime ideal of $\mathbb{Z}[\mu_N]$, we define the λ-adic versions $V^{\underline{a}}_{\lambda} = \varprojlim_m V^{\underline{a}}[\lambda^m]$. For $\underline{a} = 0$ we can and will work over $\mathbb{Z}$ rather than $\mathbb{Z}[\mu_N]$ The *geometric monodromy* of $V^{\underline{a}}_{\lambda}$ is the Zariski closure of the image of $\pi_1(T_0, t_0)$ in $Aut(V^{\underline{a}}_{\lambda,t_0})$ (for any base point t_0). Let $GL(N)^+ \subset GL(N)$ be the subgroup of elements of determinant ± 1. Say $\underline{a}$ is a *principal character* if among the a_i's there is at most one $a_i = a$ which occurs more than once (this is independent of the coset representative in (III.3.1.1)). Lemma 10.1 of [KaDw] asserts that local monodromy of $V^{\underline{a}}_{\lambda}$ at $t = \infty$ is a principal unipotent automorphism if and only if $\underline{a}$ is a principal character.

Lemma III.3.1.2 [KaDw,10.3]. *Assume $\underline{a}$ is a principal character and the repeated index a_i is 0 with multiplicity $n+1$.*

(i) In the dual case, if N is even (resp. odd) then n is odd (resp. even) and the geometric monodromy $G_{\underline{a}} = O(n)$ (resp. $Sp(n)$).

(ii) In the non-dual case, if N is even (resp. odd) then $G_{\underline{a}} = SL(n)$ (resp. $GL(N)^+$).

Definition III.3.1.3. *Hypotheses are as in III.3.1.2. Let $\mathfrak{n}$ be an ideal of $\mathbb{Z}[\mu_N]$ prime to $2N$. Let $T_{\underline{a}}[\mathfrak{n}]$ be the finite cover of T_0 parametrizing trivializations of $V^{\underline{a}}[\mathfrak{n}]$ that in the dual case preserve the pairing $<,>_{\underline{a}}$. Thus $T_{\underline{a}}[\mathfrak{n}]$ is a torsor over T_0 under the group $G_{\underline{a}}(\mathbb{Z}[\mu_N]/\mathfrak{n})$.*

The case $\underline{a} = 0$, with $N = n+1$ odd, was the basis of the potential automorphy method developed in [HST]. The following is a strengthening of the irreducibility theorem used in [HST].

Theorem III.3.1.4.

(i) [GHK, §4] *If $\underline{a} = 0$, and $N = n+1$ is odd, then $T_0[M]$ is a smooth geometrically connected curve for any M prime to $2(n+1)$. In other words, the monodromy map to $Sp(n, \mathbb{Z}/M\mathbb{Z})$ is surjective.*

(ii) [BGHT, Prop. 4.2] *More generally, suppose*

$$\underline{a} = (1, 2, \ldots, \frac{N-n-1}{2}, 0, \ldots, 0, \frac{N+n+1}{2}, \ldots, N-1)$$

with $n+1$ 0's in the middle. Let $\mathfrak{n} \subset \mathbb{Z}[\mu_N]$ be a square-free ideal and assume no two distinct prime factors of $\mathfrak{n}$ have the same residue characteristic. Then $T_{\underline{a}}[\mathfrak{n}]$ is a smooth geometrically connected curve; the monodromy map to $Sp(n, \mathbb{Z}[\mu_N]/\mathfrak{n})$ is surjective.

In [HST] (i) was proved for all M all of whose prime factors are sufficiently large. An analogue of this theorem for $T_{\underline{a}}[M]$ with monodromy $SL(n)$ is proved by Barnet-Lamb in [B-L1] and applied to potential modularity questions in [B-L2].

The authors of [HST] were led to consider the Dwork family because of the fact, highlighted by the mirror symmetry conjectures, that, for nonsingular Y_t, $H^{N-1}(Y_t)^{H_0}$ has Hodge numbers $H^{p,N-1-p}$ all equal to one, $p = 0, 1, \ldots, N-1$. This is calculated analytically, over $\mathbb{C}$, using Griffiths' theory of variation of Hodge structure of hypersurfaces, which also determines the Picard-Fuchs equation as an explicit ordinary differential equation of hypergeometric type. See §7 of [KDw], which treats more general characters of H_0 as well as more general Dwork families, and also studies the same sheaves using characteristic p methods. Let

(III.3.1.5) $$V_{t,\ell} := H^{N-1}(Y_t, \mathbb{Q}_\ell)^{H_0}; V_t[\ell] := H^{N-1}(Y_t, \mathbb{F}_\ell)^{H_0}$$

The calculation of the Hodge numbers shows that, when $t \in F$, the natural representation $\rho_{t,\ell}$ of Γ_F on $V_{t,\ell}$ is HT regular, and thus of the sort considered in the Generalized Fontaine-Mazur Conjecture. In [HST] the conjecture is in fact proved for such $\rho_{t,\ell}$ in many cases, using the theorems of §III.2.3. The method of potential automorphy on which this argument is illustrated in other applications in the following sections.

III.3.2. The Sato-Tate Conjecture for elliptic curves.

Let E be an elliptic curve over $\mathbb{Q}$. Its reduction modulo p is smooth for all but finitely many prime numbers p, and for such p the number of points $|E(\mathbb{F}_p)|$ is equal to $1 + p - a_p$, where a_p is an integer satisfying Hasse's estimate

$$-1 \leq \frac{a_p}{2\sqrt{p}} \leq 1.$$

The Sato-Tate Conjecture asserts that, **if E has no complex multiplication, then the numbers $\frac{a_p}{2\sqrt{p}}$ are equidistributed in the interval $[-1, 1]$ with respect to the probability measure $\frac{2}{\pi}\sqrt{1-t^2}dt$.** Serre showed in [S68] how to deduce this conjecture from a weak version of the tensor product functorialty conjecture of I.3.5, with $n = 2$, $F = \mathbb{Q}$, and $\tau = Sym^i$ the ith symmetric power representation of $GL(2)$. The generalizations of the Sato-Tate Conjecture considered in [S94] require not only tensor product functoriality but some version of the Tate Conjecture on cycle classes in ℓ-adic cohomology.

The Sato-Tate Conjecture was proved in [CHT,HST,T] for elliptic curves with non-integral j-invariant, using the versions III.2.3.1-2 of the Modularity Lifting Theorems; for the prime v' in condition (iii) one can take any valuation such that $v'(j(E)) < 0$. The appearance of III.2.3.3 and III.2.3.4 has made the restriction on the j-invariant superfluous, and the article [BGHT] concludes (Corollary 8.9) with the proof of the Sato-Tate Conjecture for general elliptic curves without complex multiplication over totally real fields; more precisely, [BGHT] concludes by explaining how the arguments of [CHT,HST,T] now apply without restriction.

Sketches of the argument have appeared in several places, including [H3]. Here is another one, with several steps removed, thanks to the results of [Ge]. Let π_E be the cuspidal automorphic representation of $GL(2)_{\mathbb{Q}}$ attached to E by [BCDT], in the standard normalization, so that $L(s, \pi_E) = L(s, E)$ and the functional equation exchanges $L(s, \tilde{E})$ and $L(2-s, E)$. We do not know that the functorial transfer of

π under the L-homomorphism attached to $\tau = Sym^{n-1}$ exists as an automorphic representation of $GL(n)_{\mathbb{Q}}$, but Hasse's estimate shows that the associated Langlands L-function $L(s, \pi_E, Sym^{n-1})$ is an Euler product that converges absolutely for $Re(s) > \frac{n+1}{2}$. In [S68] Serre derived the Sato-Tate Conjecture from the following analytic fact:

III.3.2.1 Theorem. *For all $n > 1$, $L(s, \pi_E, Sym^{n-1})$ has a meromorphic continuation to $\mathbb{C}$ that is holomorphic and* **non-vanishing** *for $Re(s) \geq \frac{n+1}{2}$ and satisfies the expected functional equation.*

This theorem is proved in [CHT,HST,T] for even n when $j(E) \notin \mathbb{Z}$; the case of odd n involves an extra step, which we describe in the following section. The first uniform proof is contained in [BGHT]. $M = \ell \cdot \ell'$ where ℓ, ℓ' are primes not dividing $2(n+1)$. Assume E is ordinary at ℓ, and let $\rho_{E,\ell}$ denote the representation of $\Gamma_{\mathbb{Q}}$ on $H^1(E, \mathbb{Q}_\ell)$, the dual of the Tate module. We choose a CM field $\mathcal{K}$ abelian over $\mathbb{Q}$ of degree n and a Hecke character χ of $\mathcal{K}$ of motivic type (type A_0, in Weil's terminology), so that the associated abelian character $\chi_{\ell'} : \Gamma_{\mathcal{K}} \to \bar{\mathbb{Q}}_{\ell'}^\times$ has the property that $I_{\ell'}(\chi) : \Gamma_{\mathbb{Q}} \to GL(n, \bar{\mathbb{Q}}_{\ell'})$ is absolutely irreducible and crystalline with Hodge-Tate weights $0, 1, \ldots, n-1$, each of multiplicity one. Thus both $\rho^n_{E,\ell} = Sym^{n-1}\rho_{E,\ell}$ and $I_{\ell'}(\chi)$ have the same Hodge-Tate weights as the representation of (III.3.1.5) for $t \in T_0(\overline{\mathbb{Q}})$. It can be arranged that

III.3.2.2. ℓ' is **ordinary** for $I_{\ell'}(\chi)$ and *split in the field of coefficients of* $I_{\ell'}(\chi)$ and the residual representation $I_{\ell'}(\chi) \pmod{\ell'}$ has big image in $GL(n, \mathbb{F}_{\ell'})$. Moreover, there is a non-degenerate bilinear pairing as in (c) of Theorem III.1.2.2[27]:

$$I_{\ell'}(\chi) \otimes I_{\ell'}(\chi) \to \mathbb{Q}_\ell(1-n).$$

Now it follows from Theorem III.3.1.4 and a variant of a theorem of Moret-Bailly [MB] that there is a totally real field F, Galois over $\mathbb{Q}$, and a point $t \in T_0(F)$, such that, as representations of Γ_F,

$$\rho_t[\ell] \xrightarrow{\sim} \rho^n_{E,\ell} \pmod{\ell} \; ; \; \rho_t[\ell'] \xrightarrow{\sim} I_{\ell'}(\chi) \pmod{\ell'}. \tag{III.3.2.3}$$

Here $\rho_t[\ell]$ denotes the natural action of Γ_F on $V_t[\ell]$. Now the construction of the automorphic induction in [AC] implies that $I_{\ell'}(\chi)$ *is an automorphic Galois representation* of the type constructed in III.1.2.2; using the Hodge-Tate weights one calculates that the coefficient system W of III.1.2.2 (ii) is trivial. By (III.3.2.3) $\rho_t[\ell']$ has an automorphic lifting. All the residual representations in (III.3.2.3) are ordinary by hypothesis, so we can apply Geraghty's theorem III.2.3.4[28] By III.3.2.2 and Theorem III.2.3.4, $\rho_{t,\ell'}$ is automorphic over F; since the $\rho_{t,\bullet}$ form a compatible system, so is $\rho_{t,\ell}$. Applying the same argument to the first isomorphism in (III.3.2.3), we find that $\rho^n_{E,\ell} \mid_{\Gamma_F}$ is attached to a cuspidal automorphic representation π^n_E of $GL(n)_F$. Since $F/\mathbb{Q}$ is Galois, an argument using Brauer's theorem (cf. [H3,§6]), together with the fact, due to Jacquet and Shalika, that the L-function of a cuspidal automorphic representation of $GL(n)$ does not vanish along the line of absolute convergence, implies that $L(s, \pi_E, Sym^{n-1})$ has the properties indicated in Theorem III.3.2.1.

[27] Complex conjugation is unnecessary because the base field is $\mathbb{Q}$.

[28] In particular, it is unnecessary to assume ℓ unramified in F; this saves us the need to choose an auxiliary elliptic curve and a third prime ℓ", as in [HST].

The argument works because the set of ordinary primes for E has Dirichlet density 1. This property is not known for modular forms of weight > 3. Moreover, the method sketched above appears to work only when the residue field of $\rho_{f,\lambda}$ is the prime field. Finally, and most seriously, if f is a cuspidal modular newform of weight k and $\rho_{f,\lambda}$ is Deligne's associated λ-adic representation for some prime $\lambda \mid \ell$ of the coefficient field $E(f)$ of f [D68], $\rho^n_{f,\lambda} := Sym^{n-1}\rho_{f,\lambda}$ has Hodge-Tate weights $0, k-1, 2(k-1), \ldots, (n-1)(k-1)$, which is not the collection of weights occurring in any $\rho_t[\ell]$ for $t \in T_0(\overline{\mathbb{Q}})$ if $k > 2$. The following section explains how these issues are resolved in [BGHT].

Remark III.3.2.3. Over function fields, the analogue of the Sato-Tate Conjecture was proved long ago by H. Yoshida, using geometric methods [Y].

III.3.3. The Sato-Tate conjecture for modular forms.

Let $f = \sum_{n>0} a_n(f)q^n$ be a cuspidal modular newform of weight k as in the last paragraph, normalized so that $a_1 = 1$. For almost all primes p, the Ramanujan estimates

$$-1 \leq \frac{a_p(f)}{2p^{\frac{k-1}{2}}} \leq 1$$

were derived by Deligne [D68] as a consequence of the Weil conjectures. Let $\pi(f)$ denote the automorphic representation of $GL(2)_{\mathbb{Q}}$ defined by f, and say f *is of CM type* if there is a quadratic imaginary field $\mathcal{K}$ such that the base change $\pi(f)_{\mathcal{K}}$ is not cuspidal; equivalently, $\pi(f)$ is invariant under twist by the quadratic character attached to $\mathcal{K}$.

Theorem III.3.3.1 [BGHT]. *Suppose f is not of CM type. Then the numbers $\frac{a_p(f)}{2p^{\frac{k-1}{2}}}$ are equidistributed in the interval $[-1,1]$ with respect to the probability measure $\frac{2}{\pi}\sqrt{1-t^2}dt$.*

This applies in particular to the Ramanujan Δ function:

Corollary III.3.3.2. *The numbers $\frac{\tau(p)}{2p^{\frac{11}{2}}}$, where $p \mapsto \tau(p)$ is the Ramanujan τ-function, are equidistributed in the interval $[-1,1]$ with respect to the probability measure $\frac{2}{\pi}\sqrt{1-t^2}dt$.*

There are many other special cases; for example Corollary 8.6 of [BGHT] applies Theorem III.3.3.1 to the number of ways a prime can be written as the sum of 12 squares.

As in the previous section, it suffices to prove the analogue of Theorem III.3.2.1, with π_E replaced by $\pi(f)$. For this it suffices to show that

Theorem III.3.3.3 [BGHT], Theorem B. *Let $E(f) = \mathbb{Q}(\{a_p(f)\})$ (a finite extension of $\mathbb{Q}$) and let λ be a prime of $E(f)$ dividing the rational prime ℓ. For all $n > 1$, the representation $\rho^n_{f,\lambda} := Sym^{n-1}\rho_{f,\lambda}$ is potentially automorphic over a totally real Galois extension F of $\mathbb{Q}$. Moreover, there are sufficiently many ℓ such that, if f is not ordinary at ℓ, then F can be assumed unramified above ℓ.*

The theorem already illustrates the important distinction between ordinary and other primes for f. We say f is (unramified and) *ordinary at ℓ* if the ℓ does not divide the level of f and if the Fourier coefficient a_ℓ is prime to ℓ; equivalently, the restriction of $\rho_{f,\lambda}$ to a decomposition group Γ_ℓ at ℓ is upper triangular. If f

is not ordinary at ℓ then there is fortunately only one other possibility for residual inertia: the restriction of $\bar{\rho}_{f,\lambda}$ to an inertia group I_ℓ at ℓ breaks up over $\bar{\mathbb{F}}_\ell$ as the sum of the two characters ω_2^{1-k} and $\omega_2^{\ell(1-k)}$, where ω_2 is the "second fundamental character" taking the inertia group onto the $\ell^2 - 1$'st roots of 1 in the field $\mathbb{F}_{\ell^2}$. We call such an ℓ *supersingular for* f.

When $k = 3$ it is known thanks to Wiles [Wi88] that the set of ordinary primes has density 1. Gee used this fact in [Gee] to apply the methods of [HST] to even-dimensional symmetric powers of $\rho_{f,\lambda}$. This is possible because Hida theory identifies $\bar{\rho}_{f,\lambda}$ as the residual representation attached to some ordinary newform of weight 2; it is thus as if the Hodge-Tate weights of $\bar{\rho}^n_{f,\lambda}$ are the same as those occurring in the Dwork family. This is roughly the strategy used in [BGHT] to obtain potential automorphy via ordinary primes. Now there is no reason for the restriction of $\rho^n_{f,\lambda}$ to an inertia group above ℓ to be isomorphic to the inertial representation at a point of the Dwork family; thus we need to be able to apply a Modularity Lifting Theorem that works over a base field in which ℓ is ramified. At ordinary primes [BGHT] uses III.2.3.4, which does not require that F be unramified above ℓ.

When $k > 3$ one has no idea about the distribution of ordinary and supersingular primes for f, so both cases have to be treated. Most of the difficulty in treating the ordinary case is contained in Geraghty's thesis. For ordinary ℓ, Theorem III.3.3.3 is a special case of the following theorem.

Theorem III.3.3.4 [BGHT], Theorem A. *Let F be a totally real field, n a positive integer, $\ell > 2n$ a rational prime. Let $r : \Gamma_F \to GL(n, \mathcal{O})$ be a continuous representation satisfying conditions (a) and (b) of Theorem III.1.2.2. Suppose $< \bullet, \bullet >$ is a perfect pairing on $Frac(\mathcal{O})^n$, $\mu : \Gamma_F \to \mathcal{O}^\times$ is a character, and $\varepsilon \in \pm 1$ such that*

$$< r(g)x, r(g)y >= \mu(g) < x, y >; \; < x, y >= \varepsilon < y, x >$$

and for every real prime v of F $\mu(c_v) = \varepsilon$ where c_v is a complex conjugation. Suppose moreover that r is (nearly) ordinary at all primes of F dividing ℓ and the residual representation $\bar{r}$ satisfies condition (IM) of III.2.3, and indeed

$$[\bar{F}^{\ker ad\bar{r}}(\mu_\ell) : \bar{F}^{\ker ad\bar{r}}] > 2.$$

Then there is a Galois totally real extension F'/F such that $r\,|_{\Gamma_{F'}}$ is automorphic.

If n is even, $\varepsilon = -1$, μ is the $(1-n)$th power of the cyclotomic character, and the residue field of $\mathcal{O}$ is $\mathbb{F}_\ell$, then this is almost a direct application of [HST] and III.2.3.4, as in the previous section. For general residue field k, one needs Theorem III.3.1.4(ii), with N chosen so that $\mathbb{Z}[\zeta_N]$ contains a prime ideal $\lambda \mid \ell$ such that $\mathbb{Z}[\zeta_N]/\lambda \xrightarrow{\sim} k$, and an appropriate λ' dividing ℓ' as in (III.3.2.3). We say no more about the ordinary case, except to make the following observation. Let ρ_ℓ be a compatible system of ℓ-adic representations that is "sufficiently irreducible" to satisfy the criteria of [SnW]. Suppose moreover the ρ_ℓ to be Hodge-Tate regular. If we knew that the set of ordinary primes had positive density in every generalized congruence class, then III.3.3.4 would likely suffice to prove a potential version of the Generalized Fontaine-Mazur Conjecture, or rather its version for totally real fields, for any ρ that fits in a compatible system of this type.

To treat the supersingular case one applies Guerberoff's theorem III.2.3.3. This only applies over base fields in which ℓ is unramified. An obvious problem is that

$$\rho^n_{f,\lambda}\mid_{I_\ell} \xrightarrow{\sim} \oplus_{i=0}^{n-1}\, \omega_2^{i-n+1-i\ell}(k-1) = \oplus_{i=0}^{n-1}\omega_2^{1-n+i(1-\ell)}(k-1) \qquad \text{(III.3.3.5)}$$

which does not look like the inertial representation at any point of the Dwork family. However, if in III.3.1.4(ii) we replace n by $n' = n(k-1)$ and take $\ell \equiv -1 \pmod N$, then, up to a cyclotomic twist, the inertia representation of the local system $V^{\underline{a}}_\lambda$ of III.3.1.4(ii) at the Fermat point $t = 0$ is equivalent to

$$\oplus_{i=0}^{n'-1}\omega_2^{1-n'+i(1-\ell)}. \qquad \text{(III.3.3.6)}$$

One applies a trick previously used in [H1]: find a Hecke character θ' of an abelian CM field M' of degree $2(k-1)$ such that, if $r' = I(\theta'_\ell) : \Gamma_{\mathbb{Q}} \to GL(2(k-1), \bar{\mathbb{Q}}_\ell)$ is the corresponding induced Galois representation, then, provided n' is even,

(a) The inertia representation at ℓ of $r'' = \rho^n_{f,\lambda} \otimes r'$ is equivalent to (III.3.3.6).
(b) The residual image $\bar{r''}(\Gamma_{\mathbb{Q}})$ satisfies a strengthened version ("$(k-1)$-big") of condition (IM) of III.2.3.
(c) r'' is defined over $\mathbb{Q}_\ell$.
(d) r'' is symplectic with multiplier the $(1-n')$th power of the cyclotomic character.

If n' is odd, one takes M' of degree $4(k-1)$ and chooses θ' appropriately to arrive at the same result. These conditions, which are technical but not especially difficult to satisfy, guarantee (using III.3.1.4(ii) and [MB]) that there is a totally real Galois extension $F/\mathbb{Q}$ and a point $t \in T_0(F)$ satisfying the analogue of III.3.2.3:

$$\rho_t[\ell] \xrightarrow{\sim} r'' \pmod{\ell}\ ;\ \rho_t[\ell'] \xrightarrow{\sim} I_{\ell'}(\chi) \pmod{\ell'}. \qquad \text{(III.3.3.7)}$$

Moreover, (a) above implies that we can take F *unramified at* ℓ (and even at ℓ'). Using III.2.3.3, we then conclude as in the previous section that $r'' = \rho^n_{f,\lambda} \otimes r'$ becomes automorphic over F. One then shows using base change arguments that this implies that $\rho^n_{f,\lambda}$ is also automorphic over F, completing the proof of III.3.3.3.

Hilbert modular forms. At this point my original intention was to explain what was needed to generalize the approach of [BGHT] to prove the Sato-Tate Conjecture for Hilbert modular forms. This is no longer appropriate; after completing the first draft of the present survey, I received a copy of a new article by Barnet-Lamb, Gee, and Geraghty [B-LGG] in which the Sato-Tate Conjecture is proved for Hilbert modular forms. In its general structure, the proof follows the model of [BGHT], but [B-LGG] introduces a new automorphy lifting technique, likely to apply in a variety of new situations,[29] that allows treatment of base fields in which ℓ is ramified, even when the representation is not ordinary at ℓ.

III.4 Prospects

It was mentioned in the introduction that the two experimental "tests" of Langlands conjectures launched by Langlands in the 1970s, both centered on the phenomenon of endoscopy, appear to be nearing completion. In particular, it appears

[29] My informants indicate that new applications are likely to be announced before the present article goes to press.

we are close to having constructed all compatible families of Galois representations, as well as all cases of functoriality, accessible by means of endoscopy. Langlands' article [L04] argues that, in order to prove global functorial transfer for general L-homomorphisms $\iota : {}^LH \to {}^LG$, one might begin by considering a discrete automorphic representation π of G as a contribution to the spectral side of the trace formula for G, and seek the H to which it might be functorially attached by applying the trace formula to test functions t_H designed to pick out H. Just as the subgroup $\iota({}^LH) \subset {}^LG$ is determined by the space of its invariants in representations ρ of ${}^LG$, so a π that arise from functorial transfer from a given H is expected to reveal its origins by the multiplicities $m_\pi(\rho)$ of certain poles at $s = 1$ of the corresponding Langlands L-function $L(s, \pi, \rho)$. One is led to look for a test function t_H with the property that $tr\pi(t_H) = m_\pi(\rho)$ – that is the information on the spectral side – and then to compare the geometric side of the trace formula for G, applied to t_H, with the trace formula for H.

This program is described in [L04] and continued in [L05], where Langlands makes it quite clear that even the simplest cases raise difficult questions in analytic number theory ("problèmes analytiques graves et non résolus"), as well as a whole new array of combinatorial problems. The analytic issues, at least, are more tractable when one works over function fields, as in the second part of [L05], rather than number fields, and I have heard that Langlands is working with Frenkel and Ngô to find a geometric approach to functoriality that applies to situations beyond those covered by endoscopy. More information will have to await future reports. In the meantime, I refer the reader to the recent preprint [Di] of Dieulefait, which uses the methods of automorphic lifting, combined with congruences in compatible systems, to establish many cases of non-endoscopic functorial lifting from elliptic modular forms to Hilbert modular forms over non-solvable extensions.

References

[A88] J. Arthur, The invariant trace formula II, *J. Am. Math. Soc*, **1** (1988) 501-554.

[A89] J. Arthur, The L^2 Lefschetz numbers of Hecke operators *Invent. Math*, **97** (1989) 257-290.

[A02] J. Arthur, A stable trace formula. I. General expansions, *J. Inst. Math. Jussieu*, **1** (2002) 175-277.

[A05] Arthur, J., An introduction to the trace formula, in J. Arthur, D. Ellwood, R. Kottwitz, eds, *Harmonic Analysis, the trace formula, and Shimura varieties*, *Clay Mathematics Proceedings*, **Vol. 4** (2005) 1-263.

[A08] J. Arthur, lectures at Banff International Research Station, August 2008, videos accessible at http://www.birs.ca/publications/workshop_videos/2008.html

[A10] J. Arthur, The Endoscopic Classification of Representations: Orthogonal and Symplectic Groups (book to appear).

[AC] J. Arthur and L. Clozel, *Simple Algebras, Base Change, and the Advanced Theory of the Trace Formula*, *Ann. of Math. Studies*, **120** (1989).

[B-L1] T. Barnet-Lamb, Potential automorphy for certain Galois representations to $GL(2n)$, manuscript (2008).

[B-L2] T. Barnet-Lamb, On the potential automorphy of certain odd-dimensional Galois representations, manuscript (2009).

[B-LGG] T. Barnet-Lamb, T. Gee, and D. Geraghty, The Sato-Tate conjecture for Hilbert modular forms, manuscript (2009).

[BGHT] T. Barnet-Lamb, D. Geraghty, M. Harris, and R.Taylor, A family of Calabi-Yau varieties and potential automorphy II, manuscript 2009.

[BC] J. Bellaïche and G. Chenevier, Families of Galois representations and Selmer groups, *Astérisque*, **324** (2009).

[BR92] D. Blasius and J. Rogawski, Tate classes and arithmetic quotients of the two-ball, in [LR], 421–444.

[BR93] D. Blasius and J. Rogawski, Motives for Hilbert modular forms, *Invent. Math.*, **114** (1993), 55–87.

[BW] A. Borel and N. Wallach, *Continuous cohomology, discrete subgroups, and representations of reductive groups*, Second edition, *Mathematical Surveys and Monographs*, **67**, Providence, RI: American Mathematical Society (1999).

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

[BG] K. Buzzard and T. Gee, The conjectural connections between automorphic representations and Galois representations (preliminary version fall 2009).

[COGP] P. Candelas, X. de la Ossa, P. Green, L. Parkes, An exactly soluble superconformal theory from a mirror pair of Calabi-Yau manifolds, *Nuclear Physics B*, **359** 29 July 1991, Pages 21-74.

[Book1] L. Clozel, M. Harris, J.-P. Labesse, and B. C. Ngô, eds., *The stable trace formula, Shimura varieties, and arithmetic applications, Book 1* (in preparation; preliminary versions available at http://fa.institut.math.jussieu.fr/node/44).

[Book2] L. Fargues and M. Harris, eds., *The stable trace formula, Shimura varieties, and arithmetic applications, Book 2* (in preparation).

[BT] F. Bruhat and J. Tits, Groupes reductifs sur un corps local. I. Donnees radicielles valuees *Publ. math. IHES*, **41** (1972), 5–251; II. Schémas en groupes. Existence dune donnée radicielle valuée, *Publ. Math. IHES*, **60** (1984), 5–184.

[Bump] D. Bump, Spectral theory and the trace formula, in J. Bernstein and S. Gelbart, eds. *An Introduction to the Langlands Program*, Boston: Birkhäuser (2003) 153-196.

[Ca] H. Carayol, Formes modulaires et représentations galoisiennes à valeurs dans un anneau local complet. p-adic monodromy and the Birch and Swinnerton-Dyer conjecture, *Contemp. Math.*, **165** (1994) 213-237.

[Ch] P.-H. Chaudouard, Sur le transfert lisse des intégrales orbitales d'après Waldspurger, chapter in [Book1].

[Che] G. Chenevier, Une application des variétés de Hecke des groupes unitaires, Book 2.

[CH] G. Chenevier and M. Harris, Construction of automorphic Galois representations, II, chapter in [Book2].

[C90] L. Clozel, The fundamental lemma for stable base change. *Duke Math. J.* , **61** (1990) 255–302.

[C91] L. Clozel, Représentations Galoisiennes associées aux représentations automorphes autoduales de GL(n), *Publ. Math. I.H.E.S.*, **73**, 97-145 (1991).

[CHL1] L. Clozel, M. Harris, J.-P. Labesse, Endoscopic transfer, chapter in [Book1].

[CHL2] L. Clozel, M. Harris, J.-P. Labesse, Construction of automorphic Galois representations, I, chapter in [Book1].

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

[CL] L. Clozel and J.-P. Labesse, Orbital integrals and distributions, to appear in Shahidi volume.

[CHLo] R. Cluckers, T. Hales, and F. Loeser, Transfer Principle for the Fundamental Lemma, chapter in [Book1].

[CSS] G. Cornell, J. Silverman, and G. Stevens, *Modular Forms and Fermat's Last Theorem*, New York: Springer-Verlag (1998).

[DDT] Darmon, H., F. Diamond, and R. L. Taylor: Fermat's Last Theorem, in *Current Developments in Mathematics, 1995*, Boston: International Press (1994).

[DT] J.-F. Dat and N. D. Tuan, Lemme Fondamental pour les algèbres de Lie d'après Ngô Bao-Châu, chapter in [Book1]

[D68] P. Deligne,Formes modulaires et reprsentations l-adiques. Séminaire Bourbaki, 11 (1968-1969), Exp. No. 355.

[D71] P. Deligne, Travaux de Shimura, Séminaire Bourbaki, Exposé 389, *Lecture Notes in Math.*, **244** (1971).

[D] F. Diamond, The Taylor-Wiles construction and multiplicity one, *Invent. Math.*, **128**, (1997), 379–391.

[Di] L.V. Dieulefait, Langlands Base Change for GL(2), manuscript (2009).

[FaMa] L. Fargues, E. Mantovan, *Variétés de Shimura, espaces de Rapoport-Zink et correspondances de Langlands locales*, *Astérisque*, **291** (2004).

[FM] J.-M. Fontaine and B. Mazur, Geometric Galois Representations, in *Elliptic curves, modular forms, and Fermats last theorem (Hong Kong 1993)*, Internat. Press, Cambridge MA, (1995) 41-78.

[F] K. Fujiwara, Deformation rings and Hecke algebras in the totally real case, version 2.0 (preprint 1999).

[Gee] T. Gee, The Sato-Tate conjecture for modular forms of weight 3 (manuscript 2009).

[Ge] D. Geraghty, Modularity lifting theorems for ordinary Galois representations, preprint(2009).

[GJ] S. Gelbart, H. Jacquet, A relation between automorphic representations of $GL(2)$ and $GL(3)$. *Ann. Sci. ENS* **11** (1978) 471–542.

[GHK] R. Guralnick, M. Harris, and N.M. Katz, Automorphic realization of Galois representations, manuscript (2008).

[GKM] M. Goresky, R. Kottwitz, and R. MacPherson, Discrete series characters and the Lefschetz formula for Hecke operators *Duke Math. J.*, **89** (1997), 477-554.

[Gu] L. Guerberoff, Modularity lifting theorems for Galois representations of unitary type, Thesis, Université Paris 7 (in preparation).

[Hn05] T. Haines, Introduction to Shimura varieties with bad reduction of parahoric type, in J. Arthur, D. Ellwood, R. Kottwitz, eds, *Harmonic Analysis, the trace formula, and Shimura varieties*, *Clay Mathematics Proceedings*, **Vol. 4** (2005) 583-642.

[Ha93] T. Hales, A simple definition of transfer factors for unramified groups, in *Representation theory of groups and algebras* , *Contemp. Math.*, **145**, Amer. Math. Soc., Providence, RI, (1993) 109–134.

[Ha95] T. Hales, On the fundamental lemma for standard endoscopy: reduction to unit elements, *Canad. J. Math.* , **47** (1995) 974–994.

[H1] M. Harris, Potential automorphy of odd-dimensional symmetric powers of elliptic curves, and applications, to appear in Y. Tschinkel and Yu. Zarhin, eds., *Algebra, Arithmetic, and Geometry: Volume II: In Honor Of Y.I. Manin*, Boston: Birkhäuser.

[H2] M. Harris , An introduction to the stable trace formula, chapter in [Book1].

[H3] M. Harris , Galois representations, automorphic forms, and the Sato-Tate Conjecture, to appear in *Proceedings of the Clay Research Conferences, 2007-8.*

[HL] M. Harris and J.-P. Labesse, Conditional base change for unitary groups, *Asian J. Mathematics*, **8** (2004) 653-684.

[HST] Harris, M., Shepherd-Barron, N, Taylor, R., A family of Calabi-Yau varieties and potential automorphy, *Ann. of Math.*, in press.

[HT] Harris, M. and R. Taylor, *The geometry and cohomology of some simple Shimura varieties*, *Ann. of Math. Studies*, **151** (2001).

[He] G. Henniart, Une preuve simple des conjectures de Langlands pour $GL(n)$ sur un corps p-adique, *Invent. Math.***139** (2000) 439-455.

[I] Y. Ihara, Hecke polynomials as congruence ζ-functions in elliptic modular case, *Ann. of Math.*, **85**, (1967) 267-295.

[JS] Jacquet, H. and J. Shalika, On Euler products and the classification of automorphic representations. I. *Amer. J. Math.* **103** (1981) , 499–558; II. *Amer. J. Math.* **103** (1981) , 777–815.

[Ka96] N. Katz, *Rigid local systems*, *Ann. of Math. Studies*, **139**, (1996).

[KaDw] N. Katz, Another look at the Dwork family, to appear in Y. Tschinkel and Yu. Zarhin, eds., *Algebra, Arithmetic, and Geometry: Volume II: In Honor Of Y.I. Manin*, Boston: Birkhäuser.

[KLR] C. Khare, M. Larsen, and R. Ramakrishna, Transcendental ℓ-adic Galois representations *Math. Res. Letters*, **12** (2005) 685-700.

[Ki] H. Kim , Functoriality for the exterior square of $GL(4)$ and the symmetric fourth of $GL(2)$ *JAMS* **16** (2003) 139–183.

[KS] H. Kim and F. Shahidi, Functorial products for $GL_2 \times GL_3$ and the symmetric cube for GL_2, *Ann. of Math.* **155** (2002) 837–893.

[KiMF] M. Kisin, Moduli of finite flat group schemes and modularity, *Ann. of Math.*, **170** (2009) 1085–1180.

[KiFM] M. Kisin, The Fontaine-Mazur conjecture for $GL(2)$, *J. Amer. Math. Soc.*, **22** (2009) 641–690.

[Ki2] M. Kisin, Modularity of 2-adic Barsotti-Tate representations, *Invent. Math.* **178** (2009), 587–634.

[K84] R. Kottwitz , Stable trace formula: cuspidal tempered terms, *Duke Math. J.* 51 (1984), 611-650.

[K86] R. Kottwitz, Stable trace formula: elliptic singular terms, *Math. Ann.* 275 (1986), 365-399.

[K86a] R. Kottwitz, Base change for unit elements of Hecke algebras. *Compositio Math.*, **60** (1986) 237-250.

[K88] R. Kottwitz, Tamagawa Numbers, *Ann. of Math.*, **127** (1988) 629-646 .

[K90] R. Kottwitz, Shimura varieties and λ-adic representations, in *Automorphic Forms, Shimura Varieties, and L-functions*, New York: Academic Press (1990), Vol. 1, 161-210.

[K92a] R. Kottwitz, On the λ-adic representations associated to some simple Shimura varieties, *Invent. Math.*, **108** (1992) 653-665.

[K92b] R. Kottwitz, Points on some Shimura varieties over finite fields, *JAMS*, **5** (1992) 373-444.

[K99] R. Kottwitz, Transfer factors for Lie algebras, *Representation Theory*, **3** (1999) 127-138.

[K05] R. Kottwitz, Harmonic analysis on reductive p-adic groups and Lie algebras, in J. Arthur, D. Ellwood, R. Kottwitz, eds, *Harmonic Analysis, the trace formula, and Shimura varieties*, *Clay Mathematics Proceedings*, **Vol. 4** (2005), 393-522.

[KSh] R. Kottwitz, D. Shelstad , *Foundations of twisted endoscopy*, *Astérisque*, **255** (1999).

[Lab90] J.-P. Labesse, Fonctions élémentaires et lemme fondamental pour le changement de base stable, *Duke Math. J.*, **61** (1990), no. 2, 519–530.

[Lab99] J.-P. Labesse, *Cohomologie, stabilisation et changement de base*, *Astérisque*, **257** (1999) 1-116.

[Lab01] J.-P. Labesse, Nombres de Tamagawa des groupes réductifs quasi-connexes *Manuscripta Math.*, **104** (2001) 407–430 .

[Lab04] J.-P. Labesse. Stable twisted trace formula: elliptic terms. *J. Inst. Math. Jussieu*, **3** (2004), 473–530.

[Lab09] J.-P. Labesse, Changement de base CM et séries discrètes, chapter in [Book1], with appendix by L. Clozel.

[LL] J.-P. Labesse and R. P. Langlands, L-indistinguishability for $SL(2)$, *Can. J. Math.*, **31** (1979) 726-785.

[Lf] L. Lafforgue, Chtoucas de Drinfeld, formule des traces d'Arthur-Selberg et correspondance de Langlands, *Proceedings of the International Congress of Mathematicians*, Vol. I (Beijing, 2002), 383–400

[L70] Langlands, R. P. Problems in the theory of automorphic forms. *Lectures in modern analysis and applications, III*, *Lect. Notes in Math.*, **170**, (1970), 18–61.

[L76] R. P. Langlands, Some contemporary problems with origins in the Jugendtraum, *Proc. Symp. Pure Math.*, **28**, (1976) 401-418.

[L79] R. P. Langlands, Automorphic representations, motives, and Shimura varieties. Ein Märchen, *Proc. Symp. Pure Math.*, **33**, Part 2, (1979) 205-246.

[L80] R. P. Langlands, Base change for $GL(2)$, *Ann. of Math. Studies*, **96** (1980).

[L83] R. P. Langlands, *Les débuts d'une formule des traces stables*, *Publications de l'Université Paris 7*, **13** (1983).

[L04] R. P. Langlands, Beyond Endoscopy, in H. Hida, D. Ramakrishnan, F. Shahidi, eds., *Contributions to Automorphic Forms, Geometry, and Number Theory: A Volume in Honor of Joseph Shalika*, Johns Hopkins University Press (2004) 611-698.

[L05] R. P. Langlands, Un nouveau point de repère dans la théorie des formes automorphes, at http://publications.ias.edu/rpl/.

[LRa] R. P. Langlands and M. Rapoport, Shimuravarietäten und Gerben, *J. reine angew. Math.* (1987).

[LR] R. P. Langlands and D. Ramakrishnan, eds., *The zeta functions of Picard modular surfaces*, Université de Montréal, Centre de Recherches Mathé matiques, Montreal (1992).

[LS87] R. P. Langlands, and D. Shelstad, On the definition of transfer factors, *Math. Ann.*, **278** (1987), 219-271.

[LS90] R. P. Langlands and D. Shelstad, Descent for transfer factors, *The Grothendieck Festschrift, Vol. II*, *Progr. Math.*, **87**, Birkhäuser Boston, Boston, MA, (1990) 485–563.

[LN] G. Laumon and B. C. Ngô, Le lemme fondamental pour les groupes unitaires, *Ann. of Math.* **168** (2008) 477–573.

[Mz89] B. Mazur, Deforming Galois representations, in Y. Ihara, K. Ribet, J.-P. Serre, eds., *Galois groups over* $\mathbb{Q}$, New York: Springer-Verlag (1989), 385-437.

[Mz08] B. Mazur, Finding meaning in error terms, *Bull. AMS*, **45** (2008) 185-228.

[Ml05] J. S. Milne, Introduction to Shimura varieties, in J. Arthur, D. Ellwood, R. Kottwitz, eds, *Harmonic Analysis, the trace formula, and Shimura varieties*, *Clay Mathematics Proceedings*, **Vol. 4** (2005) 265-378.

[Mi] A. Minguez, Unramified representations of unitary groups, chapter in [Book1].

[M07] C. Moeglin, Classification et changement de base pour les séries discrètes des groupes unitaires p-adiques, *Pac. J. Math.*, **233** (2007), p. 159-204.

[M09] C. Moeglin, Paquets d'Arthur discrets pour un groupe classique p-adique, in D. Ginzburg, E. Lapid, and D. Soudry, eds., *Automorphic Forms and L-functions II: Local Aspects* (Proceedings of the conference in honor of S. Gelbart), *Contemporary Math.* **489** (2009) 179-258.

[M09a] C. Moeglin, Multiplicité 1 dans les paquets d'Arthur, to appear in the proceedings of the conference in honor of F. Shahidi.

[Mo] S. Morel, Étude de la cohomologie de certaines variétés de Shimura non compactes, *Ann. of Math. Studies*, in press.

[MB] L. Moret-Bailly, Groupes de Picard et problèmes de Skolem II, *Ann. Scient. Ec. Norm. Sup.* **22** (1989), 181–194.

[N06] B. C. Ngô, Fibration de Hitchin et endoscopie, *Invent. Math.*, **164** (2006) 399-453.

[N2] B. C. Ngô, Le lemme fondamental pour les algèbres de Lie, in press.

[N3] B. C. Ngô, Decomposition theorem and abelian fibration, to appear in [Book1]

[Rk] R. Ramakrishna, Infinitely ramified Galois representations, *Ann. of Math.*, **151** (2000), 793–815.

[Ra] D. Ramakrishnan , Modularity of the Rankin-Selberg L-series, and multiplicity one for $SL(2)$. , *Ann. of Math.*, **152** (2000), 45-111.

[Rp] M. Rapoport, A guide to the reduction modulo p of Shimura varieties. *Automorphic forms. I*, *Astérisque*, **298** (2005) 271–318.

[R] Rogawski, J., *Automorphic Representations of Unitary Groups in Three Variables*, *Ann. of Math. Studies*, **123** (1990).

[RZ] M. Rapoport and T. Zink, *Period spaces for p-divisible groups*, *Ann. of Math. Studies*, **141** (1996).

[S68] J.-P. Serre, *Abelian ℓ-adic representations and elliptic curves*, New York: Benjamin (1968).

[S94] J.-P. Serre, Propriétés conjecturales des groupes de Galois motiviques et des représentations ℓ-adiques, in *Motives*, *Proc. Symp. Pure Math.*, **55**, Part 1 (1994) 377-400.

[Shel79] D. Shelstad, Characters and inner forms of a quasi-split group over $\mathbb{R}$, *Compositio Math.*, **39** (1979) 11–45.

[Shel82] D. Shelstad, L-indistinguishability for real groups, *Math. Ann.* , **259** (1982), 385–430.

[Sh66] G. Shimura, Moduli of abelian varieties and number theory, in *Algebraic Groups and Discontinuous Subgroups*, *Proc. Symp. Pure Math.*, **9** 312–332.

[Sh70] G. Shimura, On arithmetic automorphic functions. *Actes du Congrès International des Mathématiciens* (Nice, 1970), Tome 2, 343–348.

[Shin] S.-W. Shin, Galois representations arising from some compact Shimura varieties, *Ann. of Math.*, in press.

[SW] C. Skinner and A. Wiles, Residually Reducible Representations and Modular Forms,*Publ. Math. IHES* **89** (1999), 5-126.

[SnW] A. Snowden and A. Wiles, Bigness in compatible systems (manuscript, 2008).

[So] C. M. Sorensen, A patching lemma, to appear in [Book2].

[T02] R. Taylor, Remarks on a conjecture of Fontaine and Mazur, *J. Inst Math. Jussieu*, **1** (2002), 1–19.

[T02a] R. Taylor, Galois representations, *Proceedings of ICM 2002*, volume I, 449-474.

[T06] R. Taylor, On the meromorphic continuation of degree two L-functions. *Doc. Math.* (2006) Extra Vol., 729-779.

[T08] R. Taylor, Automorphy for some ℓ-adic lifts of automorphic mod ℓ Galois representations, II. *Publ. Math. IHES*, **108** (2008) 1-181.

[TW] R. Taylor and A. Wiles, Ring-theoretic properties of certain Hecke algebras, *Annals of Math.*, **141**, 553-572 (1995).

[W97] J.-L. Waldspurger, Le lemme fondamental implique le transfert, *Compositio Math.*, **105** (1997) 153-236.

[W01] J.-L. Waldspurger, Intégrales orbitales nilpotentes et endoscopie pour les groupes classiques non ramifiés, *Astérisque*, **269** (2001).

[W06] J.-L. Waldspurger, Endoscopie et changement de caractéristique, *J. Inst. Math. Jussieu*, **5** (2006).

[W08] J.-L. Waldspurger, L'endoscopie tordue n'est pas si tordue, *Mem. AMS*, **194** (2008).

[W09] J.-L. Waldspurger, Endoscopie et changement de caractéristique : intégrales orbitales pondérées, *Ann. Inst. Fourier*, **59** (2009) 1753-1818.

[W] J.-L. Waldspurger, A propos du lemme fondamental tordu, chapter in [Book1].

[Wi88] A. Wiles, On ordinary λ-adic representations associated to modular forms, *Invent. Math.*, **94** (1988), 529–573.

[Wi95] A. Wiles, Modular elliptic curves and Fermat's last theorem, *Annals of Math.*, **142**, 443-551 (1995).

[Y] H. Yoshida, On an analogue of the Sato conjecture, *Inventiones Math.*, **19**, 261-277 (1973).